Geneseq version 3.5
Copyright (c) 1998 - 2002 Compugen Ltd.

GMatrix... protein search, using sw model

Run on... September 17, 2002, 16:29:45 ; Search time 240.42 Seconds
(without alignments)
7631.270 Million cell updates/sec

Title... Query, 434 aa

Perfect score: 879

Sequence... 1 MAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAA 434

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched... 179999 seqs, 134826294 residues

Total number of hits satisfying chosen parameters: 359999

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database... GenEmbl:*
1: gb_ba1:*
2: gb_htg1:*
3: gb_inv:*
4: gb_mam:*
5: gb_mus:*
6: gb_pat:*
7: gb_phg:*
8: gb_pln:*
9: gb_prt:*
10: gb_rod:*
11: gb_sts:*
12: gb_syn:*
13: gb_una:*
14: gb_vir:*
15: gb_vrt:*
16: em_ba1:*
17: em_bum:*
18: em_inv:*
19: em_mus:*
20: em_om:*
21: em_pat:*
22: em_phg:*
23: em_pln:*
24: em_prt:*
25: em_rod:*
26: em_sts:*
27: em_syn:*
28: em_una:*
29: em_vir:*
30: em_htg_hum:*
31: em_htg_mus:*
32: em_htg_other:*
33: em_htg_inv:*

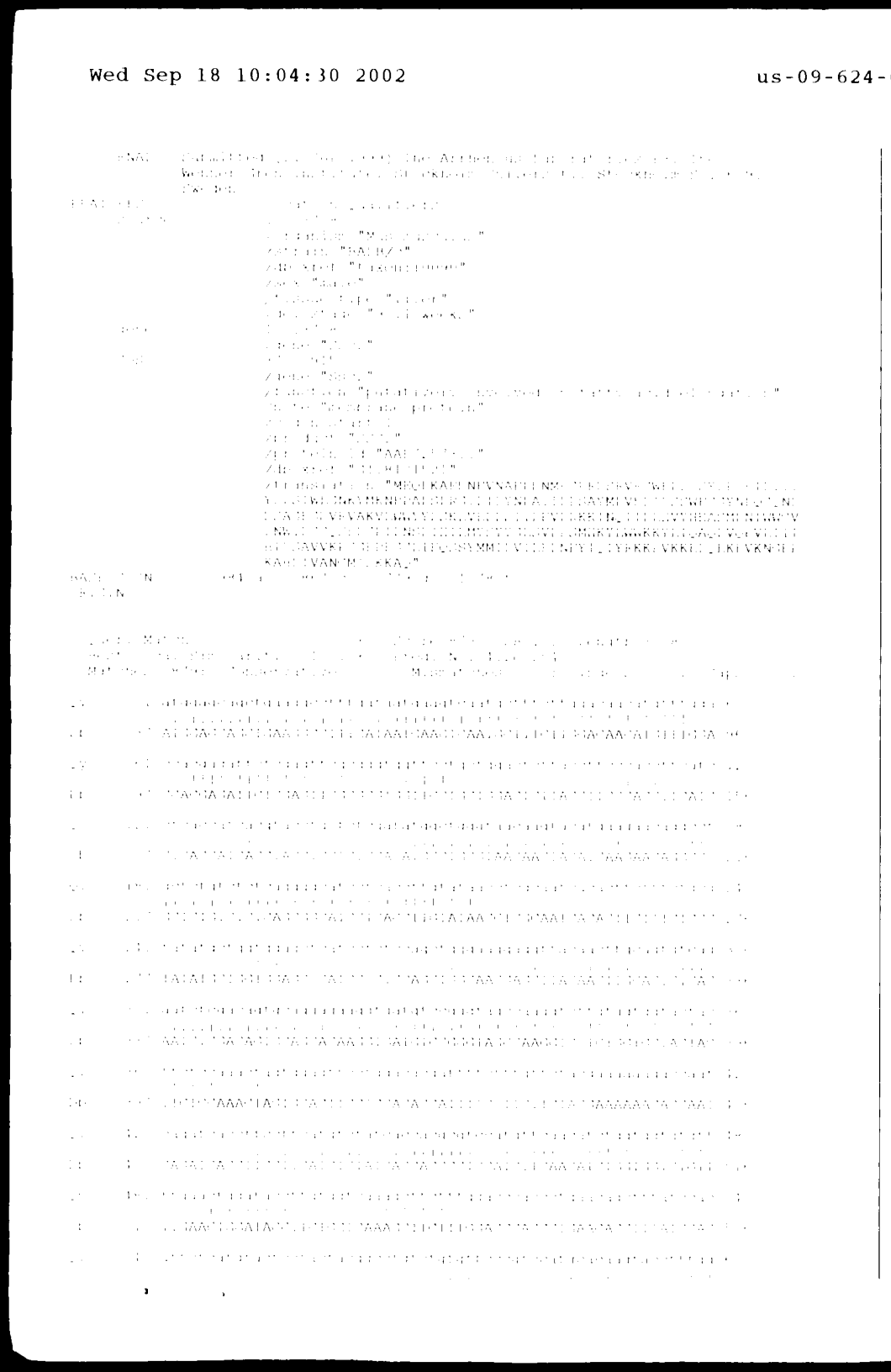
Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

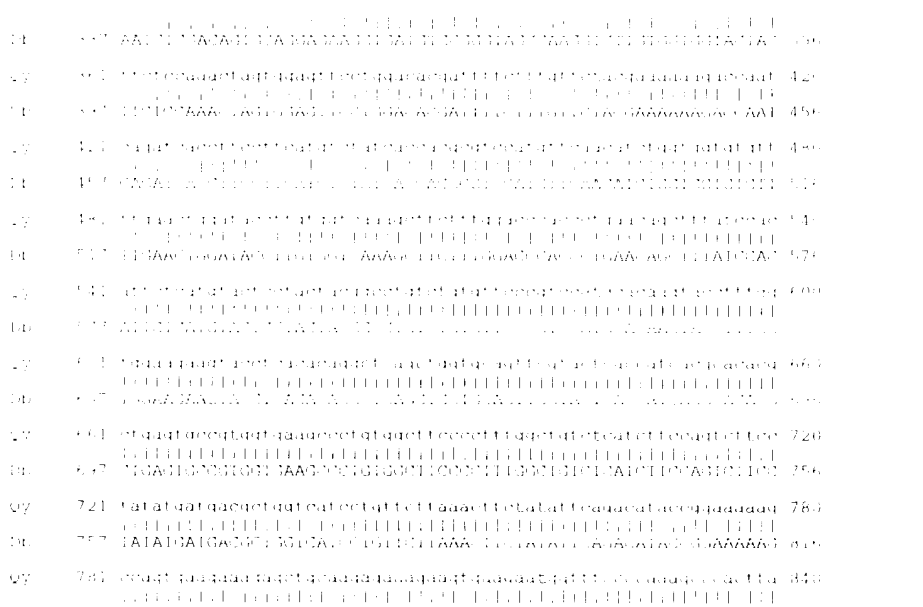
SUMMARIES

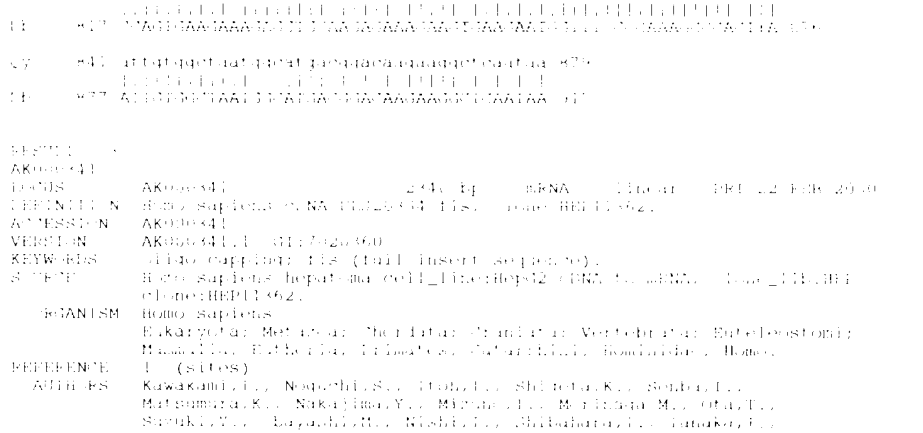
Result No.	Score	Match Length	DB	ID	Description
1	879	100.0	4304	1	AF170908 Mus musculus
2	879	100.0	4304	4	AX052774 Sequence
3	650.6	71.0	2441	9	AK000341 Homo sapi
4	548.6	59.4	2752	9	BC022611 Mus muscu
5	534.8	58.1	1397	5	AF465520 Sequential
6	512.2	57.8	7988	10	AB072985 Rattus no
7	471.4	57.2	2352	9	BX886965 Homo sapi
8	430.6	57.6	960	9	AF243981 Homo sapi
9	430.2	57.6	2892	6	Z96465 Sequence 1
10	422.2	57.6	2852	5	Z99224 Sequence 1
11	418.8	57.1	458	5	AX127718 Sequence
12	417.2	57.1	2468	10	AF277093 Mus muscu
13	417.2	57.1	3552	9	AB068106 Macaca fa
14	377.6	56.2	2279	9	AK056277 Homo sapi
15	377.6	56.2	2980	9	AF277094 Homo sapi
16	377.6	56.2	2992	9	AF057208 Homo sapi
17	375.2	14.7	114833	2	AC097662 Rattus no
18	373.6	57.8	1.05e5	9	BC036102 Homo sapi
19	369.2	57.8	43e+04	2	AL117638 Homo sapi
20	344	57.4	2e+05	14	AF196100 Rattus n
21	343.2	57.1	2500	9	AY060267 Drosophil
22	334.8	55.1	63341	2	AC018047 Drosophil
23	334.8	55.1	181950	3	AC008054 Drosophil
24	334.8	57.1	227904	3	AE003677 Drosophil
25	332.2	57.0	1830	10	BC006562 Mus muscu
26	327	14.4	3525	3	AY011182 Drosophil
27	327	14.4	5e+04	2	AC015170 Drosophil
28	327	14.4	64196	3	AC006494 Drosophil
29	327	14.4	77e+04	3	AC008197 Drosophil
30	327	14.4	223815	3	AE003739 Drosophil
31	327.8	57.7	344 13	2	AC097962 Rattus no
32	312.6	57.8	50e+04	2	AC015170 Drosophil
33	312.6	57.8	164194	3	AC006494 Drosophil
34	312.6	57.8	223815	3	AE003739 Drosophil
35	310.2	57.5	1684	3	AY060447 Drosophil
36	308.4	12.3	1471	9	AF151846 Homo sapi
37	308.8	12.2	2508	9	AK027256 Homo sapi
38	305.8	12.1	1457	9	AK001653 Homo sapi
39	305.8	12.1	1486	9	BC000618 Homo sapi
40	304.2	13.9	469	5	AX052771 Sequence
41	304.2	13.9	1471	9	AF336793 Homo sapi
42	300.6	11.4	1537	10	BC006735 Mus muscu
43	300.6	11.4	1466	10	AF170907 Mus muscu
44	300.4	11.4	1488	5	AX052769 Sequence
45	300	11.1	967	8	AF206662 Mortierel

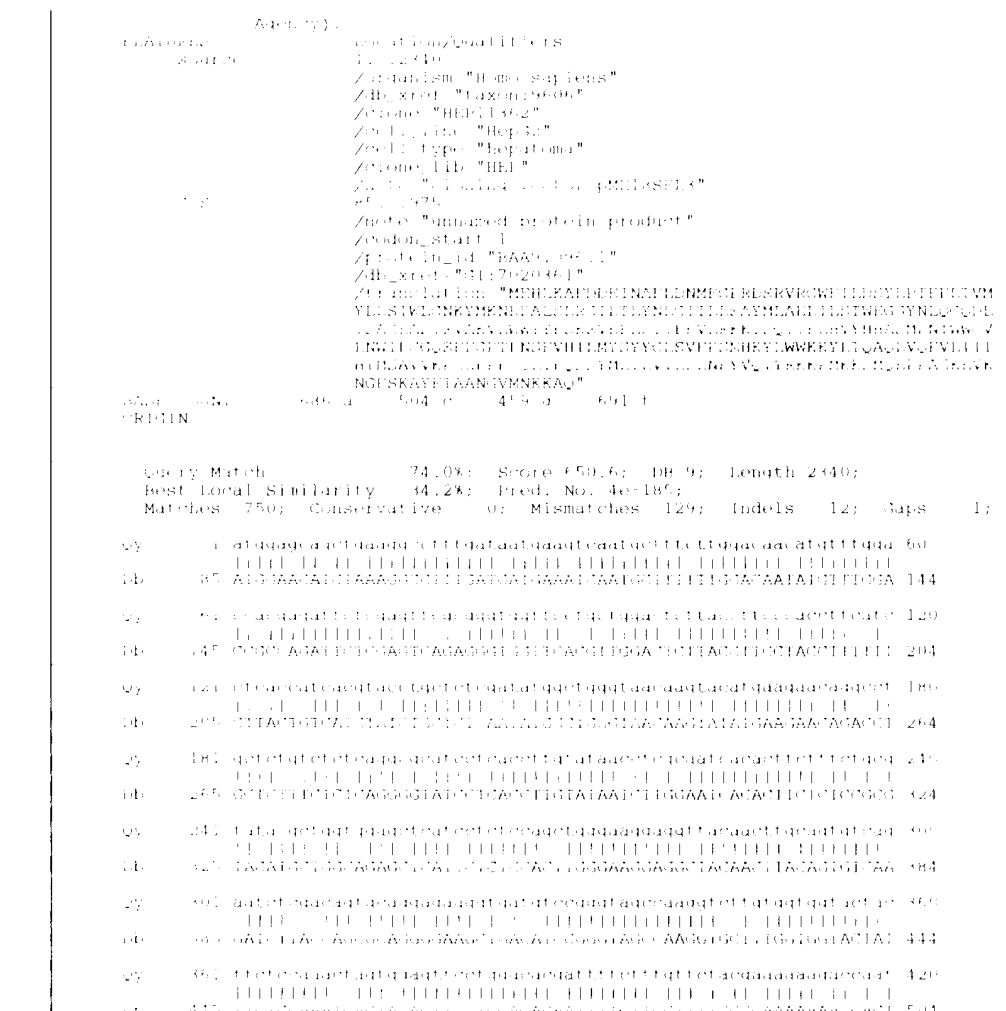
ALIGNMENTS

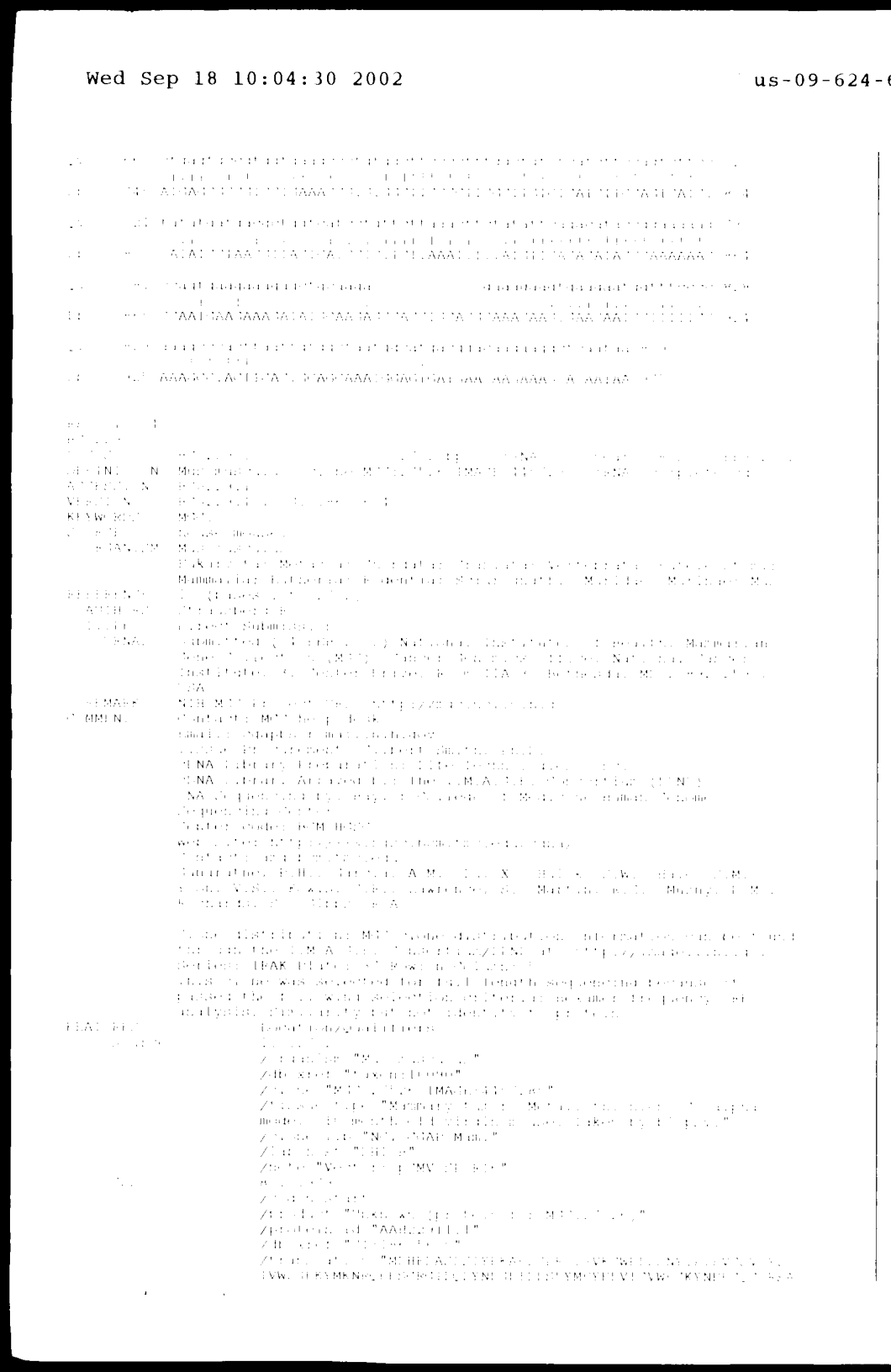
RESULT 1
AF170908
LOCUS AF170908 4304 bp mRNA linear BCT 06-MAY-2000
DEFINITION Mus musculus SSY2 (SSY2) mRNA, complete cds.
ACCESSION AF170908
VERSION AF170908.1 GI:8162615
KEYWORDS .
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 4308)
AUTHORS Ivrdik,F., Westerberg,L., Silve,S., Asadi,A., Jakobsson,A.,
Cantor,B., Naim,S. and Jacobsson,A.
TITLE Role of a new mammalian gene family in the biosynthesis of very long chain fatty acids and sphingolipids
JOURNAL J. Cell Biol. 149 (4), 707-718 (2000)
MEDLINE 20254178
PUBMED 10791985
REFERENCE 2 (bases 1 to 4308)
AUTHORS Ivrdik,F., Westerberg,L., Silve,S., Asadi,A., Jakobsson,A.,
Naimi,S., Boisen,S. and Jacobsson,A.
TITLE Direct Submission



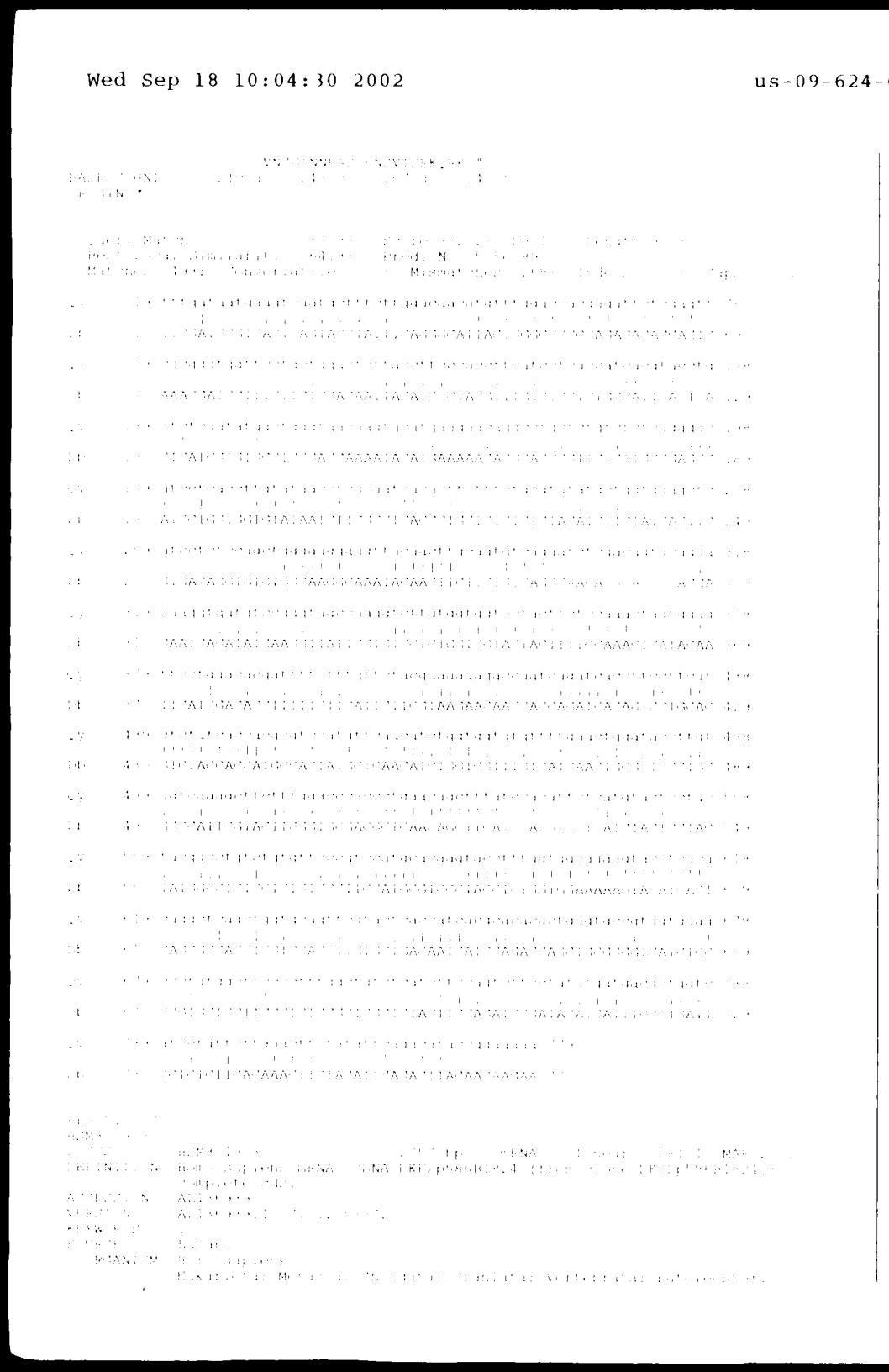


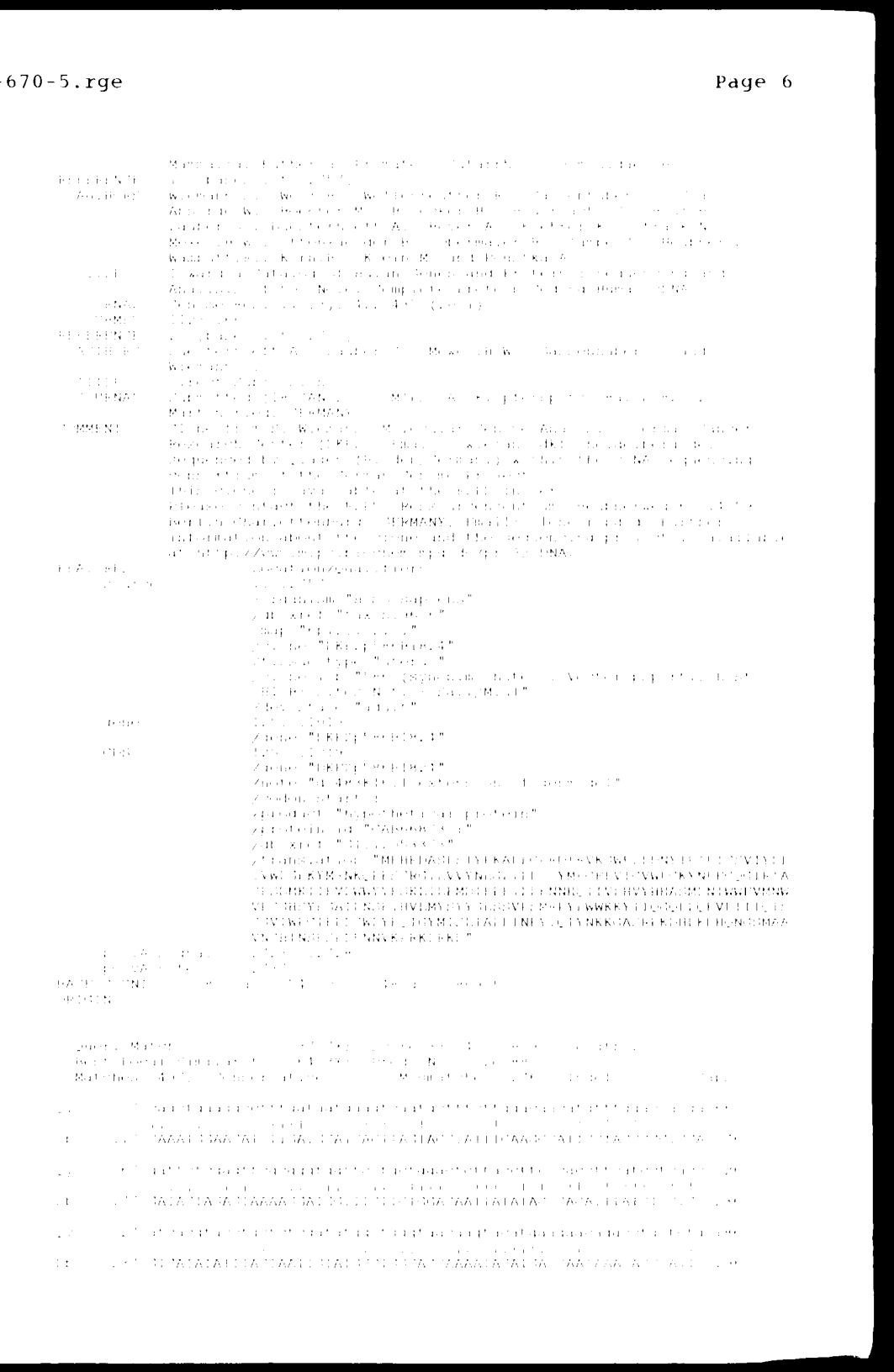






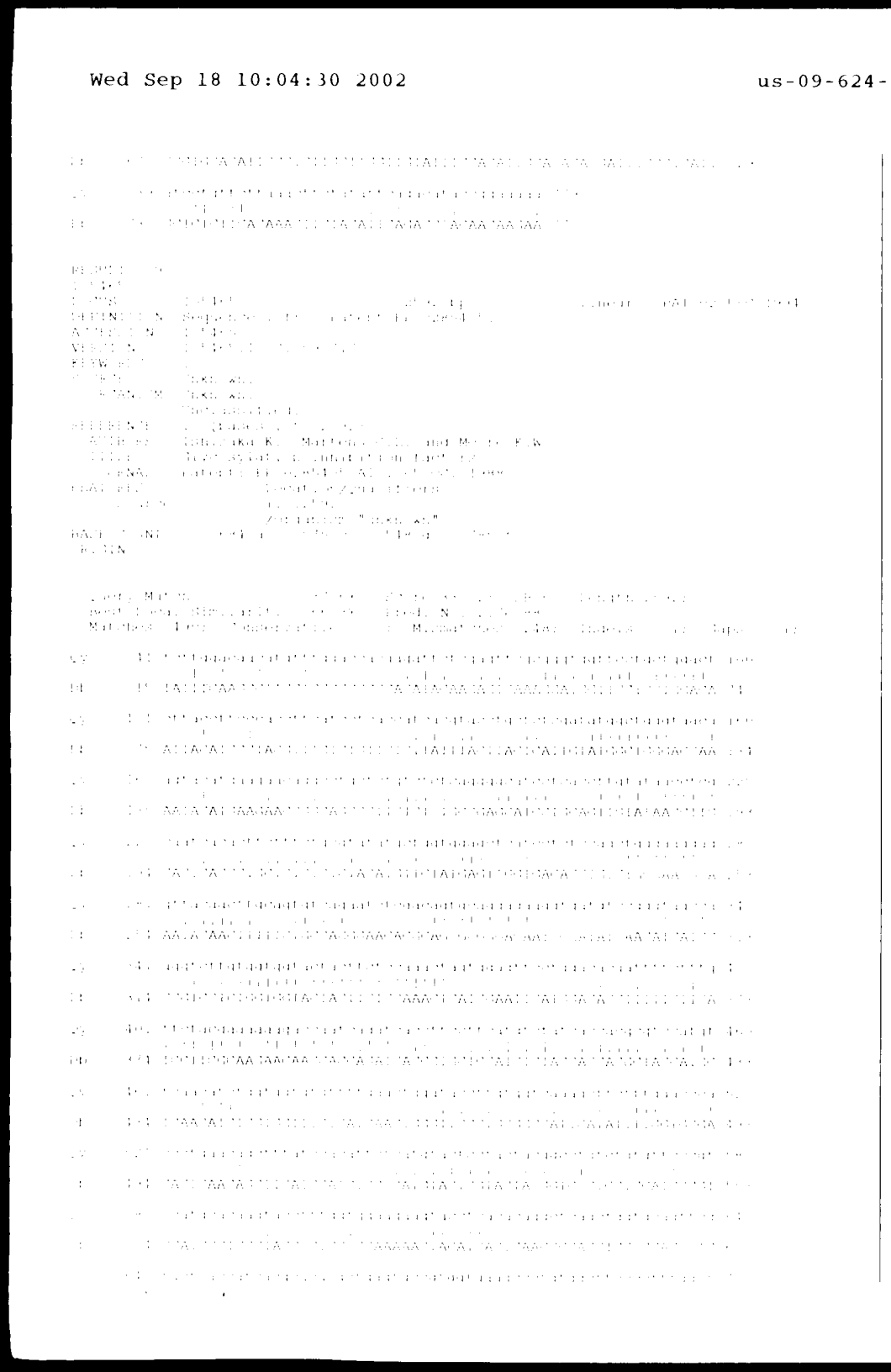
```
SOURCE      [illegible]
  ORGANISM  Scophthalmus maximus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
            Pleuronectoidei; Scophthalmidae; Scophthalmus.
REFERENCE   1  (bases 1 to 1397)
  AUTHORS   Agaba,M.K., [illegible], Dickson,C.A. and [illegible]
  TITLE     Polyunsaturated fatty acid elongase from teleost fish
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1397)
  AUTHORS   Agaba,M.K., [illegible] and [illegible]
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2002) Institute of Aquaculture, University of
            Stirling, Stirling FK9 4LA, UK
FEATURES             Location/Qualifiers
     source          1..1397
                     /organism="Scophthalmus maximus"
                     /db_xref="taxon:52904"
                     /tissue_type="[illegible]"
     gene            1..1397
                     /gene="[illegible]"
     5'UTR           1..28
                     /gene="[illegible]"
     CDS             29..[illegible]
                     /gene="[illegible]"
                     /codon_start=1
                     /product="polyunsaturated fatty acid elongase"
                     /protein_id="AA[illegible]84.1"
                     /db_xref="GI:[illegible]"
                     /translation="[illegible]
                     [illegible]
                     [illegible]
                     [illegible]
                     [illegible]
                     SNGIPSAEHMAHKKLRVD"
     3'UTR           [illegible]..1397
                     /gene="[illegible]"
BASE COUNT     [illegible] a    [illegible] c    [illegible] g    [illegible] t
ORIGIN

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length 1397;
  Best Local Similarity   64.7%;  Pred. No. [illegible];
  Matches  498;  Conservative    0;  Mismatches  272;  Indels    0;  Gaps    0;

Qy          7 [illegible] 66
Db         [illegible] 85

Qy         67 [illegible] 126
Db         [illegible] 145

Qy        127 [illegible] 186
Db        146 [illegible] 205

Qy        187 [illegible] 246
Db        [illegible] 265

Qy        [illegible] 306
Db        [illegible] 325

Qy        [illegible] 366
Db        [illegible]

Qy        [illegible] 426
Db        [illegible]

Qy        427 [illegible] 486
Db        446 [illegible] 505

Qy        487 [illegible] 546
Db        [illegible] 565

Qy        547 [illegible] 606
Db        [illegible] 625

Qy        607 [illegible] 666
Db        626 [illegible] 685

Qy        667 [illegible] 726
Db        [illegible] 745

Qy        727 [illegible] 776
Db        [illegible] 795


RESULT 6
AB071985
LOCUS       AB071985                 900 bp    mRNA    linear   ROD 16-OCT-2001
DEFINITION  Rattus norvegicus rELO1 mRNA for fatty acid elongase 1, complete
            cds.
ACCESSION   AB071985
VERSION     AB071985.1  GI:16151798
KEYWORDS    .
SOURCE      Rattus norvegicus (strain:Sprague-Dawley), male [illegible] cDNA to mRNA.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 900)
  AUTHORS   Inagaki,K., Aki,T., [illegible],Y., Kawamoto,S., Shigeta,S., [illegible],K. and
            Suzuki,O.
  TITLE     Identification and expression of a mammalian fatty acid elongase
            involved in de novo biosynthesis of C18 fatty acids
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 900)
  AUTHORS   Aki,T.
  TITLE     Direct Submission
  JOURNAL   Submitted ([illegible]) Tsunehiro Aki, Hiroshima University,
            Graduate School of Advanced Sciences of Matter; 1-3-1 Kagamiyama,
            Higashi Hiroshima, Hiroshima 739-8530, Japan
            (E-mail:akit@hiroshima-u.ac.jp, Tel:81-824-24-7755,
            Fax:81-824-24-7754)
FEATURES             Location/Qualifiers
     source          1..900
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /sex="male"
                     /tissue_type="liver"
     gene            [illegible]
                     /gene="rELO1"
     CDS             [illegible]
                     /gene="rELO1"
                     /codon_start=1
                     /product="fatty acid elongase 1"
                     /protein_id="BAB[illegible].1"
                     /db_xref="GI:16151799"
                     /translation="[illegible]
                     IVWLIPKYMKNRQPFSCRGILVVYNLGLTLLSLYMFYELVTGVWEGKYNFFCQGTRSA
                     GESDMKVIRVLWWYYF[illegible]KLIEFMDTFFFILRKNNHQITVLHVYHHATMLNIWWFVMNW
                     VPCGHSYFGATLNSFIHVLMYSYYGLSSVPSMRPYLWWKKYITQGQLVQFVLTIIQTS
                     [illegible]MI7
```

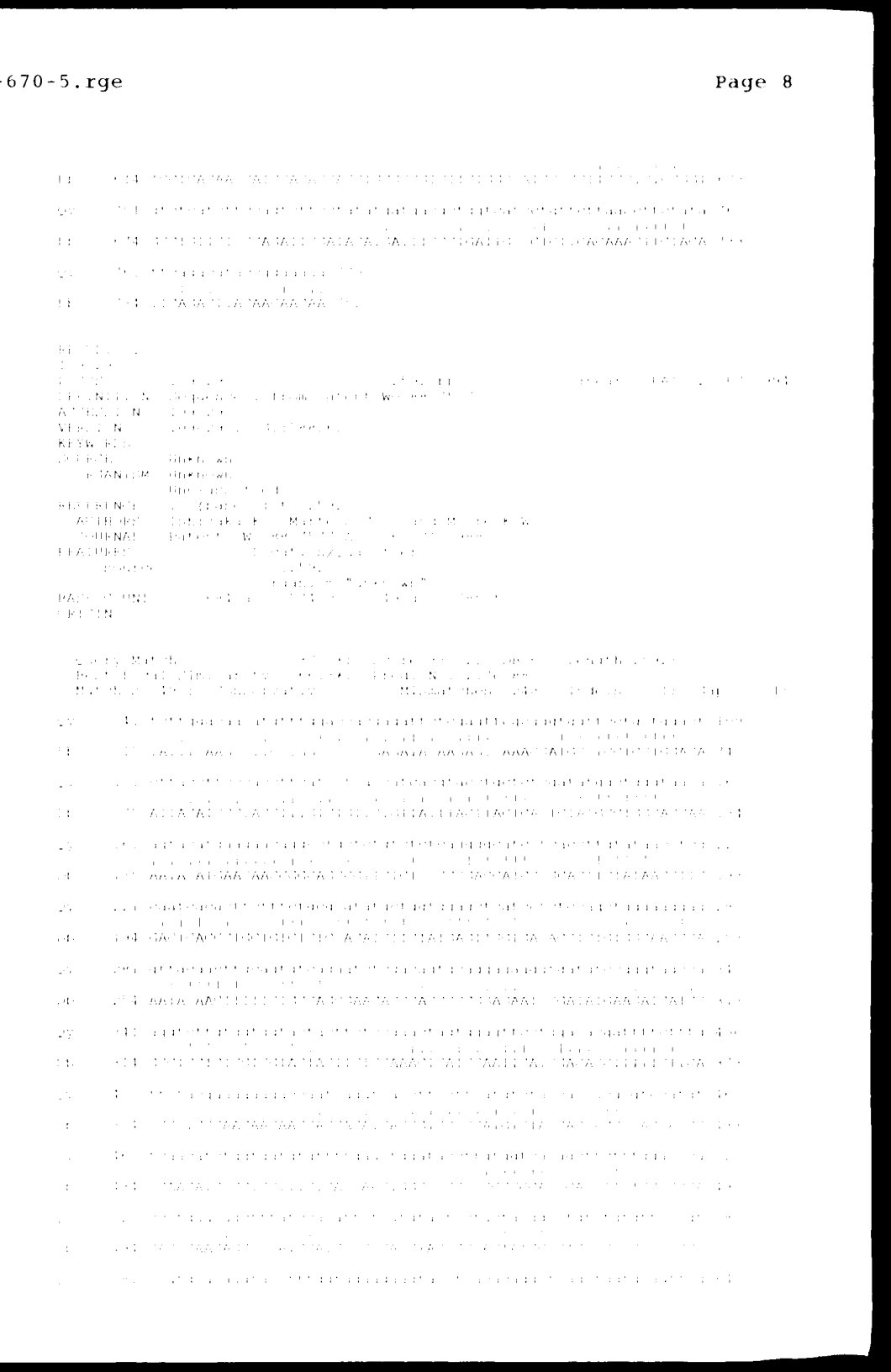


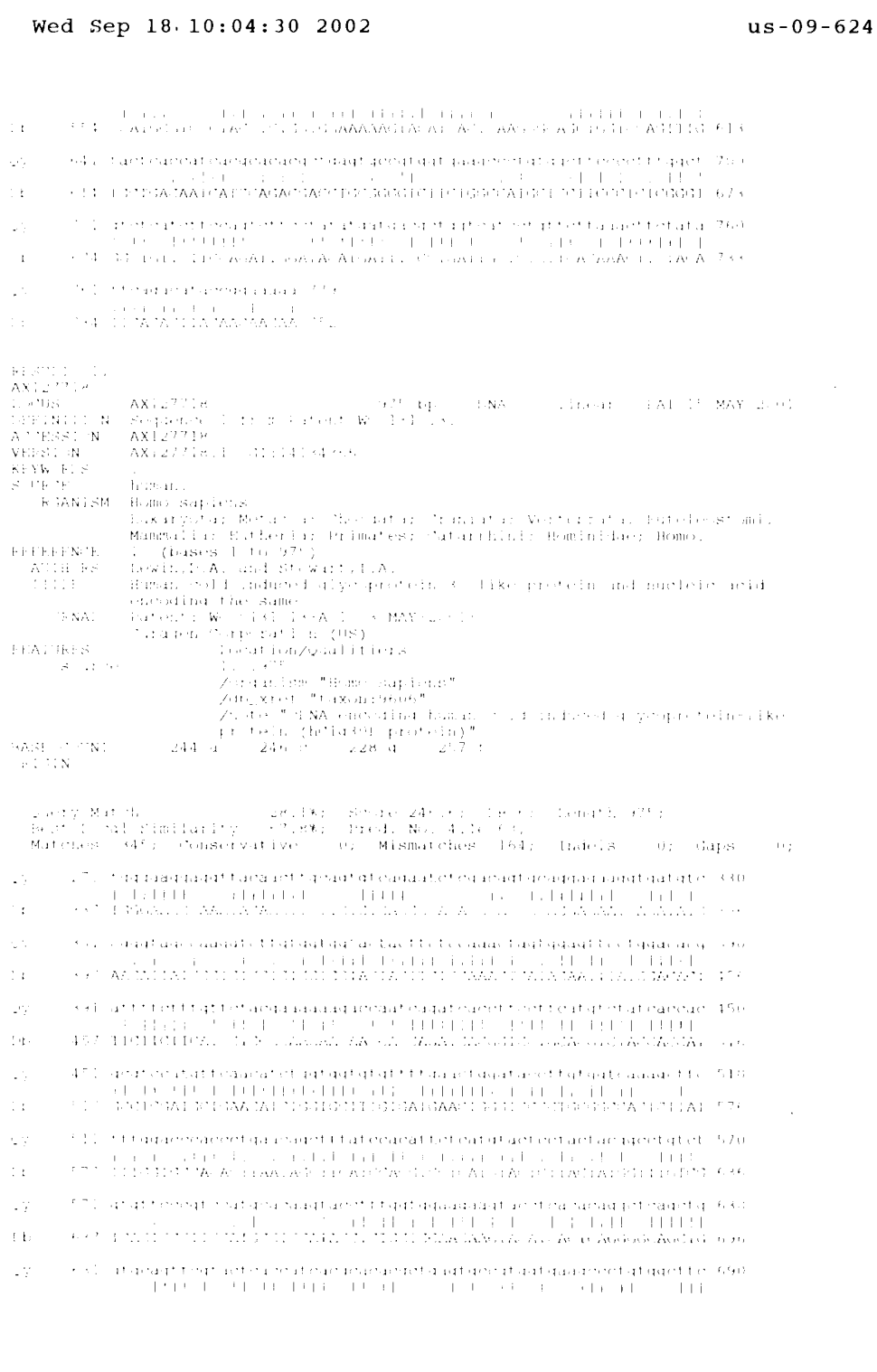


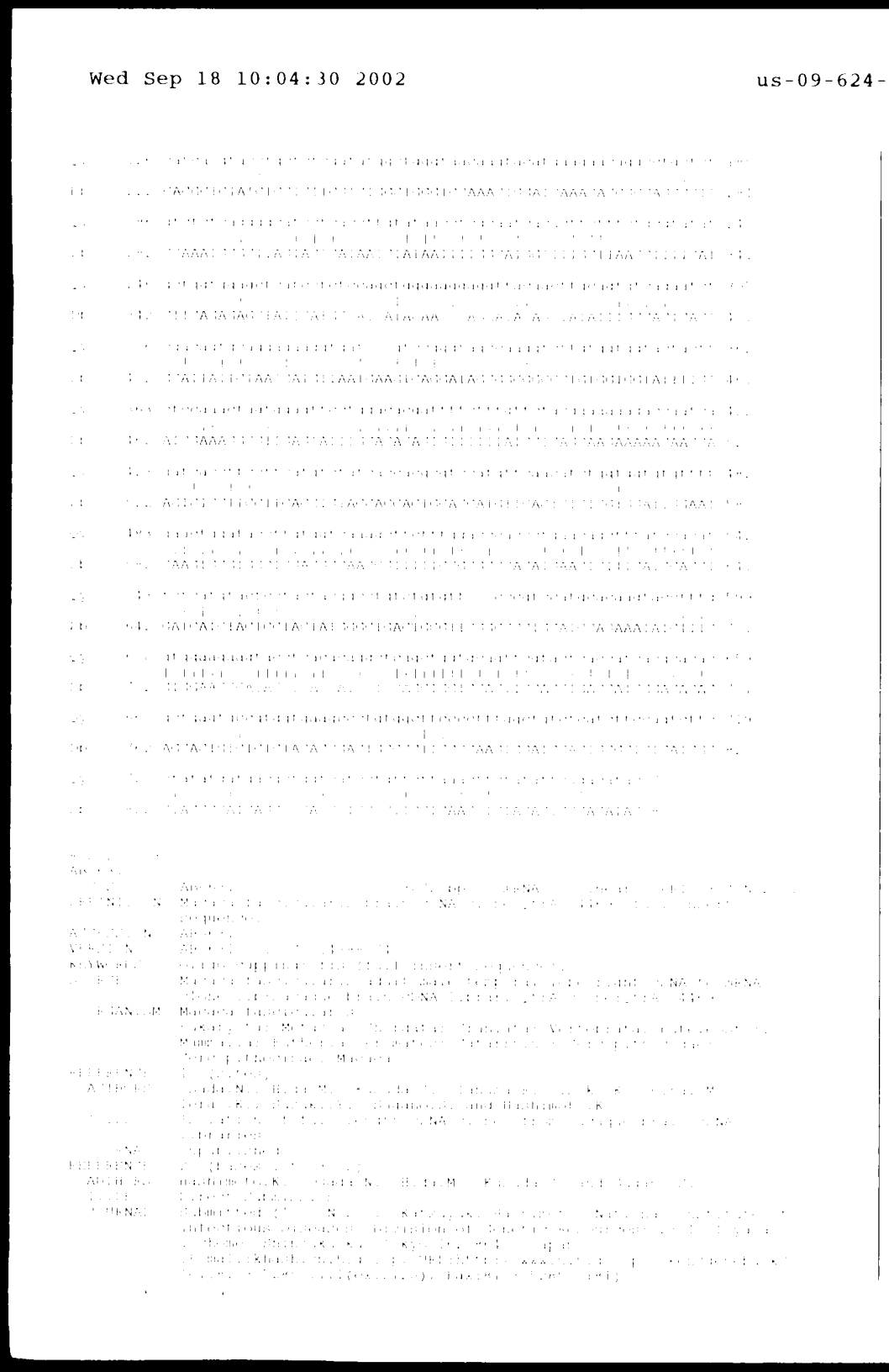
```
                     /map="6p12"
                     /note="BEC-1"
                     /function="involved in conversion of GLA to DGLA, AA to
                     adrenic acid, alpha-linolenic acid to 20:4n-3, stearidonic
                     acid to 20:4n-3, and EPA to 22:5n-3"
                     /codon_start=1
                     /product="long chain polyunsaturated fatty acid elongation
                     enzyme"
                     /protein_id="AAF70651.1"
                     /db_xref="GI:7920040"
                     /translation="MEHTVAELSYTPAELSKHSTPESWVLMVLLEKVYDLSKFVEEHPGGEEVLSEQAGKDATDDFEDVGHSTDARELMKTYIIGELHPSDRNKPSLETVKATIENNFKSNHNDNIKLVDFAIELEKNHQAINYFLQVAEMSTDDKSAQSFLGAIMKLQSQIADKKVTKENLDSLIKKGNNAHLKSVPPEISDTLEKLKAAGTKVMVVSDFNTFFIETILTAMGKDWTKYFSSIITNPSNMDKDGRLFIEIVPFDSQHECALCPPNMCKGLVIEDIQKEIGYKKMIYIGDGAFDFCPHLKLSDYDIMFPRKGYPLWDLIHNNPELAKAKITPWSDGKELLDTLLLNIIDKRESEQTLSVNKLTSEWKEGLEKEIKTVLSSLDDKELFDHGVDKSILGRVAPIQVLKLEKELVAKVLKKEVTFPTLGITGSGKSSVINAMRSLVHDEDEAAAKTGIVETTMEATAYDWPHTNIRFLDTPGHEAFTSMRARGAQVTDIVVLVVAADDGVMPQTIEAINHAKAANVPIIVAINKIDKPGANPDRVMQELSSIGLQPEEWGGDTIFVNISAKFNLNIDELLEMILLQAEVLELKANPNRMAKGTVIEAQLDKGRGAVATVLVQEGTLHKGDIVLCGLEYGRVRAMVDENGKNVKEAGPSIPVEILGLSGVPSAGDEFEVMESDLKAREIAEKRQIKLRDVQLAQQQHVSLENMFSQIQEGEVKELNIIIKADVQGSVEAIAASLLKLSTDEVKVKIIHSGVGAITESDVTLASASNAIIIGFNVRPDSNARKSAEVEGVDIRLHSIIYKLIDDVEKAMKGMLSPEYEEKVLGLAEVRQTFKISKVGTIAGCMVTDGKIARNAGIRLLRDGIVIYEGELESLKRFKDDVKEVQTGYECGIKIKNFNDIKEGDIIEAFEMIEIKRTL"
BASE COUNT 230 a 222 c 196 g 252 t
ORIGIN
```

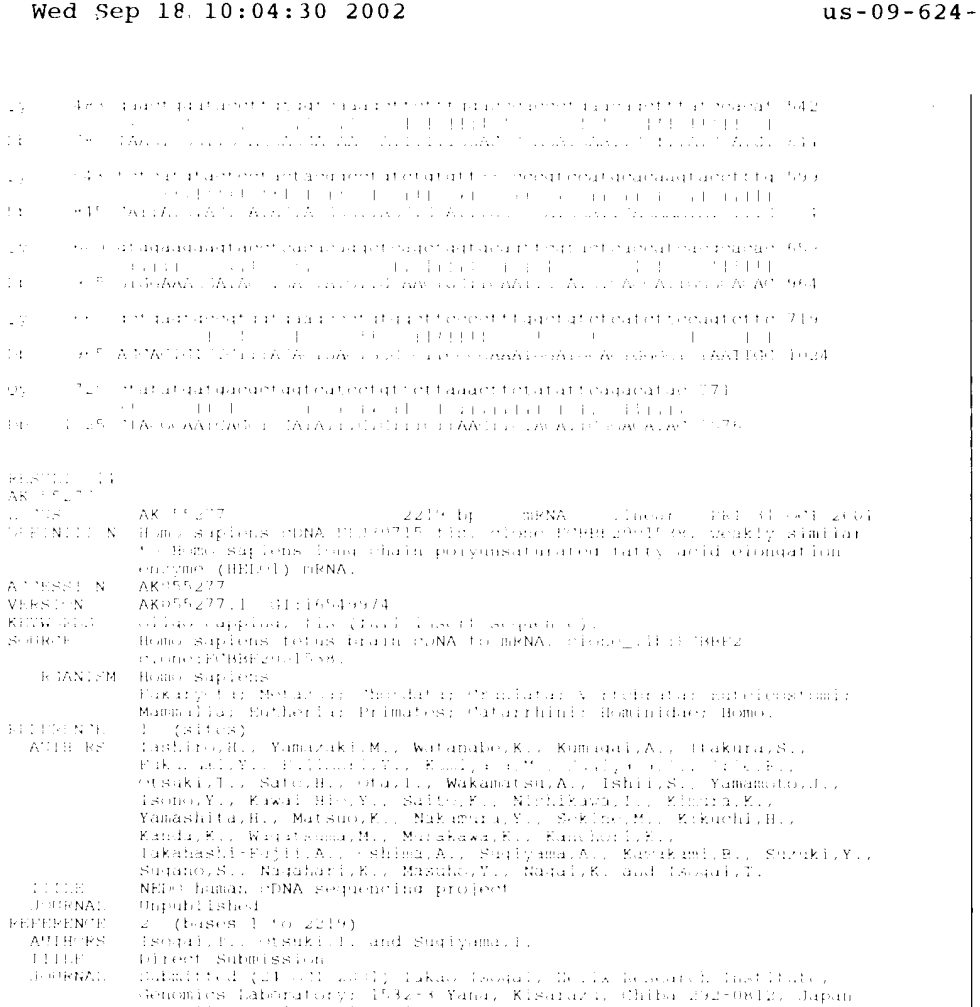
Query Match 47.4%; Score 336.4; DB 9; Length 906;
Best Local Similarity 64.7%; Pred. No. 1.7e-88;
Matches 492; Conservative 0; Mismatches 269; Indels 92; Gaps 6;



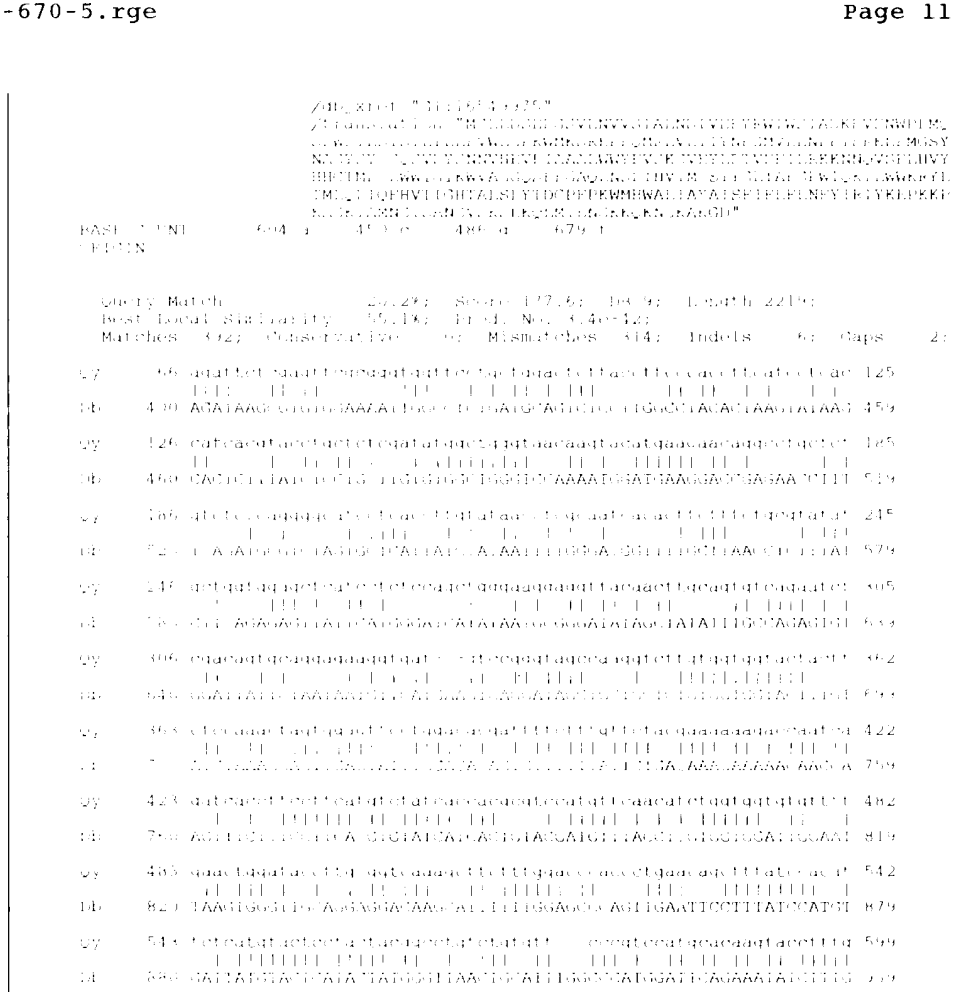




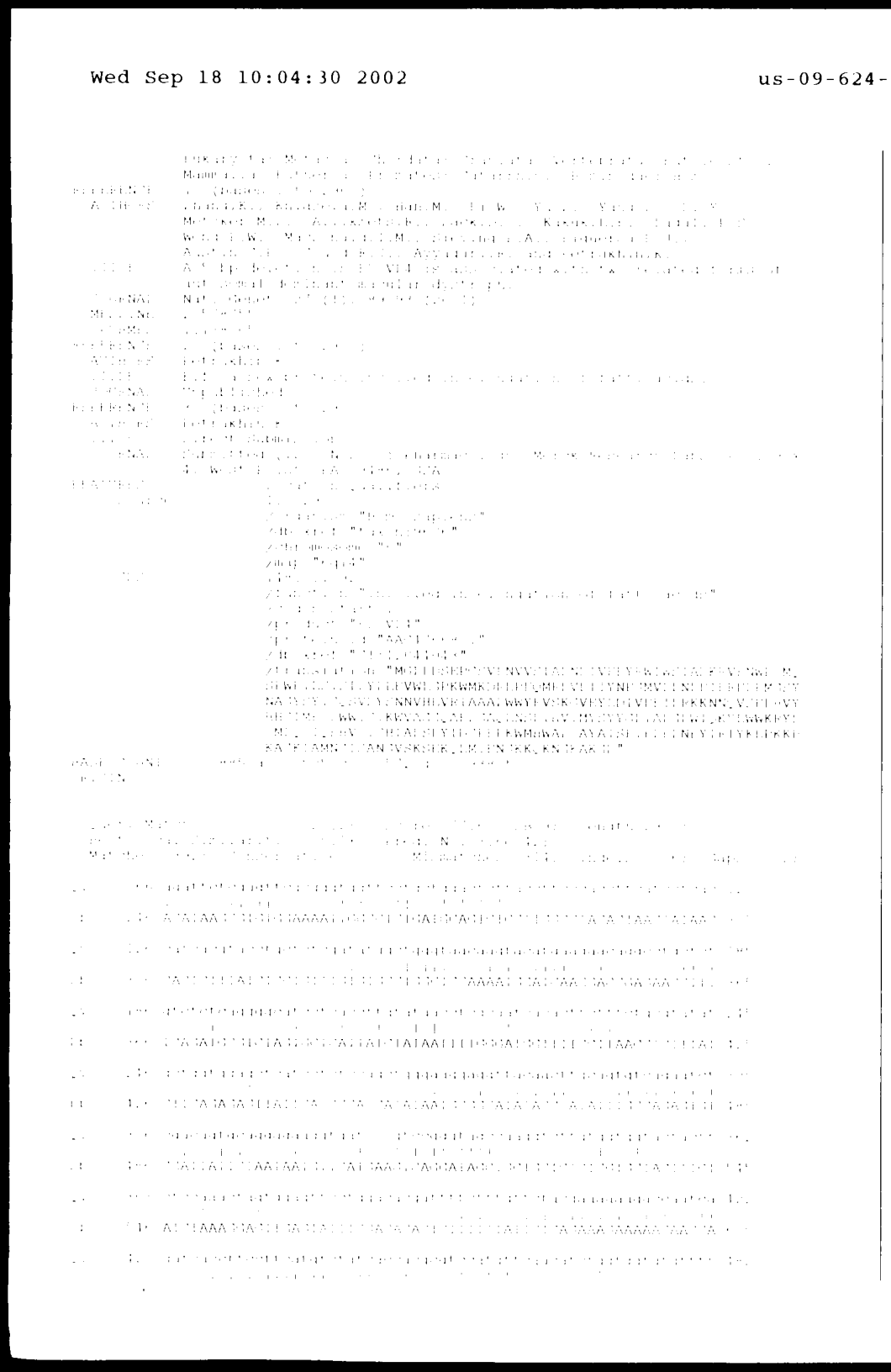




RESULT 14
AK055277
LOCUS AK055277 2215 bp mRNA linear PRI 01-OCT-2002
DEFINITION Homo sapiens cDNA FLJ30715 fis, clone FCBBF2015744, weakly similar to Homo sapiens fatty acid elongase (HELO1) mRNA.
ACCESSION AK055277
VERSION AK055277.1 GI:16549974
KEYWORDS cDNA supplement, fis (Full Insert sequence).
SOURCE Homo sapiens fetus brain cDNA to mRNA, clone:_IB:FCBBF2
clone:FCBBF2015744.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Isobe,K., Yamazaki,M., Watanabe,K., Komaoki,A., Hakura,S., Fukuda,Y., Fujimori,O., Kouji,H., Yoshida,K., Otsuki,T., Sato,H., Otui,I., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai,H.Y., Saito,F., Nishikawa,T., Kimura,K., Yamashita,H., Matsuno,K., Nakamura,Y., Soeda,M., Kikuchi,H., Kanda,K., Watakabe,M., Murakawa,K., Kamiguti,A., Takahashi-Fujii,A., Ishima,A., Shinya,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2215)
AUTHORS Isogai,T., Otsuki,T. and Sugiyama,T.
TITLE Direct Submission
JOURNAL Submitted (21-OCT-2002) Takao Isogai, Tokyo Research Institute, Genomics Laboratory; 1583-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@ri.co.jp, Tel:81 438 52 3951, Fax:81 438 52 3952)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Bolex Research Institute (BRI) (supported by Japan Key Technology Center etc.); 5' & 3' end one pass sequencing: RAB, BRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and BRI.
FEATURES Location/Qualifiers
source 1..2215
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FCBBF2015744"
/tissue_type="brain"
/clone_lib="FCBBF2"
/dev_stage="fetus"
/note="cloning vector: pME18SFL3"
CDS 362..1246
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB70965.1"



RESULT 15
AF277094
LOCUS AF277094 2990 bp mRNA linear PRI 12-JAN-200
DEFINITION Homo sapiens ElovL4 mRNA, complete cds.
ACCESSION AF277094
VERSION AF277094.1 GI:12044412
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens



Geneseq version 4.0
Copyright (c) 1998 - 2000 Compugen Ltd.

Minimum match 0% Maximum match 100%

Run on: September 17, 2002, 14:42:14 ; Search time 26+14.42 Seconds
(without alignments)
7816.770 Million cell updates/sec

Title: US-09-624-670-6
Perfect score: 906
Sequence: 1 atgtgaaattgtgagcgg.....gtaatcgcgaaatgatta 906

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1750664 seqs, 1342620439 residues

Total number of hits satisfying chosen parameters: 1750664

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: GeneSeq:*
1: gb_ba:*
2: gb_bx:*
3: gb_htg:*
4: gb_in:*
5: gb_om:*
6: gb_pat1:*
7: gb_pat2:*
8: gb_pat3:*
9: gb_ph:*
10: gb_pl:*
11: gb_pr1:*
12: gb_pr2:*
13: gb_pr3:*
14: gb_ro:*
15: gb_sts:*
16: gb_sy:*
17: em_ba:*
18: em_bx:*
19: em_htg:*
20: em_in:*
21: em_om:*
22: em_pl:*
23: em_ro:*
24: em_sts:*
25: em_sy:*
26: em_un:*
27: em_htc:*
28: em_htg_fin:*
29: em_htg_htc:*
30: em_htg_phase1:*
31: em_htg_phase2:*

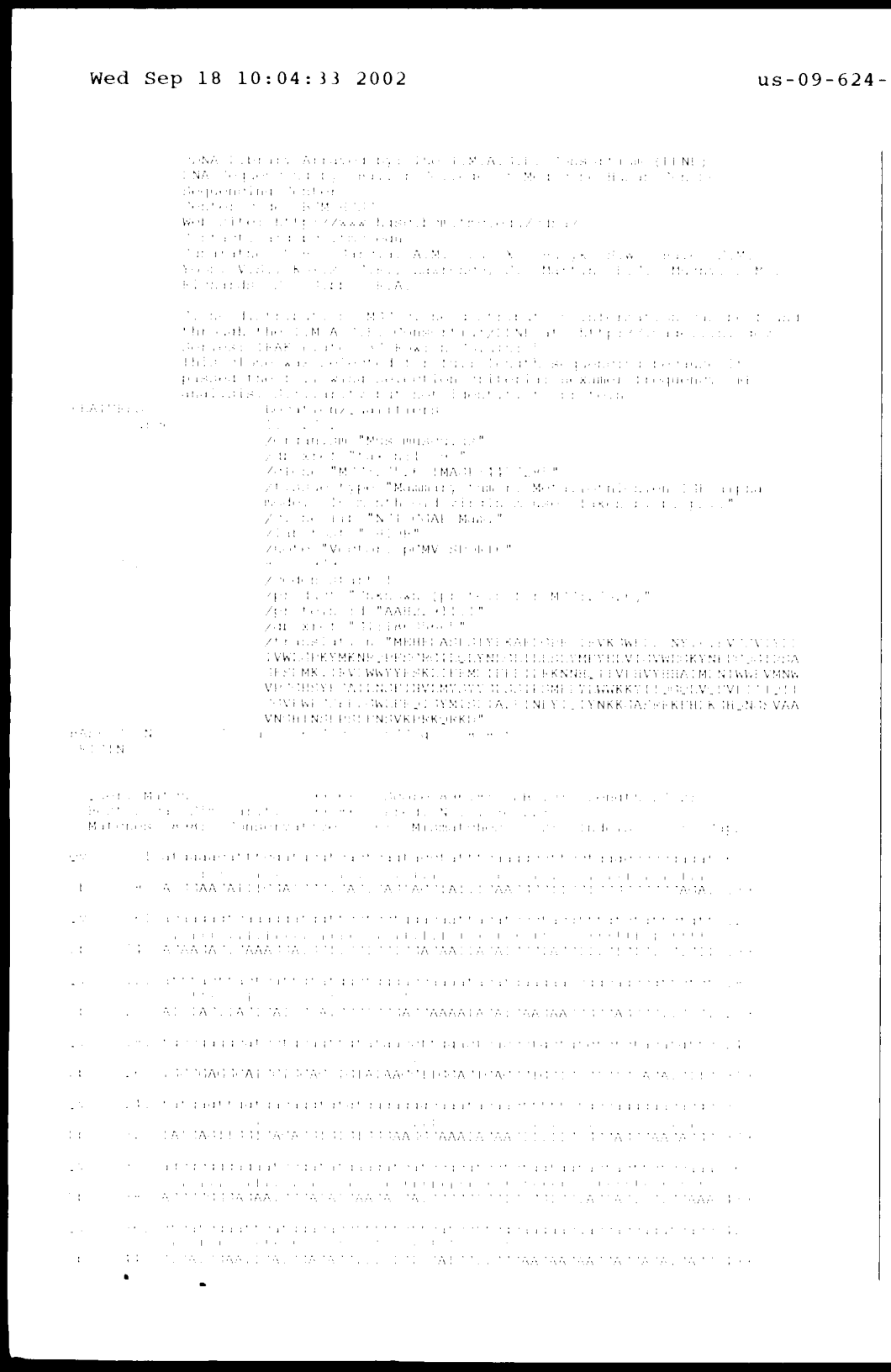
Frost No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB ID	Description
1	906.0	100.0	2752	Be022911 Be022911 Mus musculus
2	856.8	94.2	2992	6 105465 105465 Sequence 1
3	856.8	94.2	2992	6 I79029 I79029 Sequence 1
4	786.4	86.4	906	10 AB071985 AB071985 Rattus no
5	716	79.6	960	9 AF243981 AF243981 Homo sapi
6	716	79.6	2752	9 BX063589 BX063589 Homo sapi
7	584.6	66.4	975	6 AX127718 AX127718 Sequence
8	447.2	49.7	1067	5 AF465520 AF465520 Scpbtbr1
9	435	48.3	4398	10 AF178968 AF178968 Mus musculus
10	427	47.2	2752	6 AX127721 AX127721 Sequence
11	416.2	45.7	2842	9 AK000341 AK000341 Homo sapi
12	259	28.8	2066	9 BC017270 BC017270 Homo sapi
13	246.8	27.2	73665	2 AC110592 AC110592 Mus musculus
14	190	21.1	3363	10 AF277094 AF277094 Mus musculus
15	174	19.3	2209	9 AK056277 AK056277 Homo sapi
16	174	19.3	2900	9 AF277094 AF277094 Homo sapi
17	174	19.3	2992	9 AY042288 AY042288 Homo sapi
18	170.8	18.9	5072	9 AB041810 AB041810 Macaca fa
19	145.2	16.1	101270	9 HS883K16 AL034574 Human DNA
20	115.2	12.7	176655	2 AC010684 AC010684 Homo sapi
21	42	4.6	2147	5 BC002735 BC002735 Mus musculus
22	42	4.6	4408	10 AF170007 AF170007 Mus musculus
23	42	4.6	4488	9 AX052769 AX052769 Sequence
24	42	4.6	4497	9 AK056455 AK056455 Homo sapi
25	42	4.6	4498	9 BC020184 BC020184 Homo sapi
26	42	4.6	4502	9 AF151940 AF151940 Homo sapi
27	134.4	14.8	999	9 AX052771 AX052771 Sequence
28	134.4	14.8	4502	9 AF151940 AF151940 Homo sapi
29	134.4	14.8	29859	14 AF198100 AF198100 Lowlpox v
30	126.8	14.0	1884	8 AY060447 AY060447 Drosophil
31	126.8	14.0	1895	8 AY062182 AY062182 Drosophil
32	125.2	13.8	1850	10 BC056602 BC056602 Mus musculus
33	122.2	13.5	2583	8 AY062267 AY062267 Drosophil
34	120	13.2	50474	2 AC015170 AC015170 Drosophil
35	120	13.2	164193	8 AC006494 AC006494 Drosophil
36	120	13.2	174983	8 AC008197 AC008197 Drosophil
37	120	13.2	228615	8 AE003749 AE003749 Drosophil
38	117.8	13.0	151431	8 AC013647 AC013647 Drosophil
39	115.8	12.8	181950	8 AC008093 AC008093 Drosophil
40	115.8	12.8	227894	8 AE003677 AE003677 Drosophil
41	115.2	12.7	1412	9 BC009838 BC009838 Homo sapi
42	115.2	12.7	110600	2 AC024562_6 AC024562 Homo sapi
43	115.2	12.7	110600	2 AC024562_1 Continuation (2 of
44	112.2	12.5	50474	2 AC015170 AC015170 Drosophil
45	112.2	12.5	164193	8 AC006494 AC006494 Drosophil

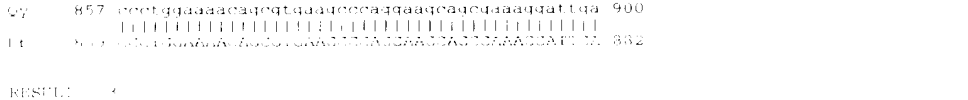
ALIGNMENTS

RESULT 1
BC022911
LOCUS BC022911 2752 bp mRNA linear ROD 07-FEB-2002
DEFINITION Mus musculus, clone MGC:27526 IMAGE:4457285, mRNA, complete cds.
ACCESSION BC022911
VERSION BC022911.1 GI:18056701
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2752)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
FEATURES NCBI staff (info@mgc.nci.nih.gov)
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.

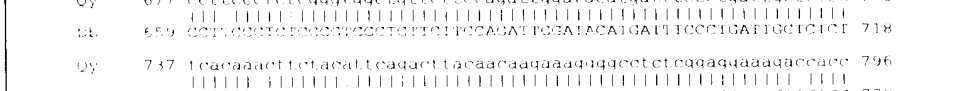


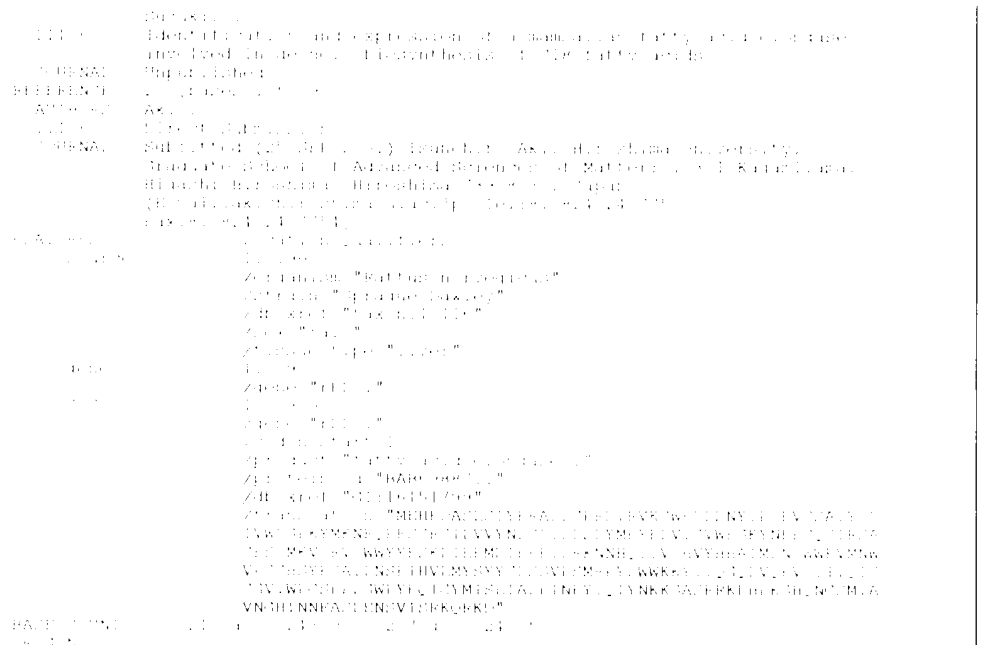
RESULT 3
I09029
LOCUS I09029 2592 bp linear PAT 02-DEC-1994
DEFINITION Sequence 2 from Patent WO 8807577.
ACCESSION I09029
VERSION I09029.1 GI:588262
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 2592)
AUTHORS Ishizuka,K., Martens,C.L. and Moore,K.W.
JOURNAL Patent: WO 8807577-A 1 06-OCT-1988;
FEATURES Location/Qualifiers
source 1..2592
/organism="Unknown"
BASE COUNT 684 a 574 c 549 g 785 t
ORIGIN

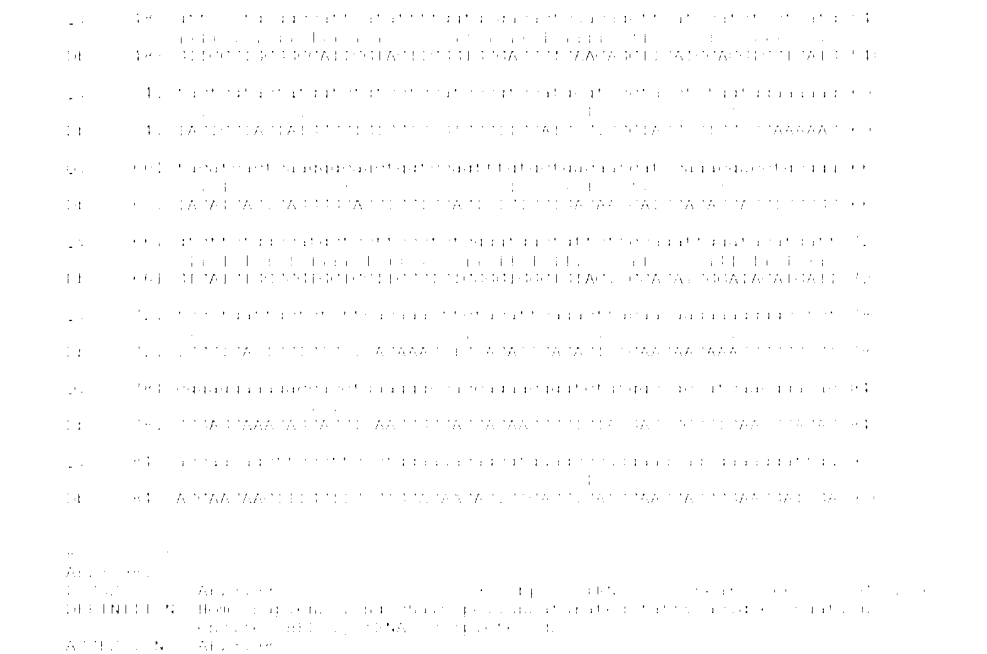
Query Match 95.2%; Score 856.8; DB 6; Length 2592;
Best Local Similarity 99.5%; Pred. No. 3.6e-215;
Matches 860; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

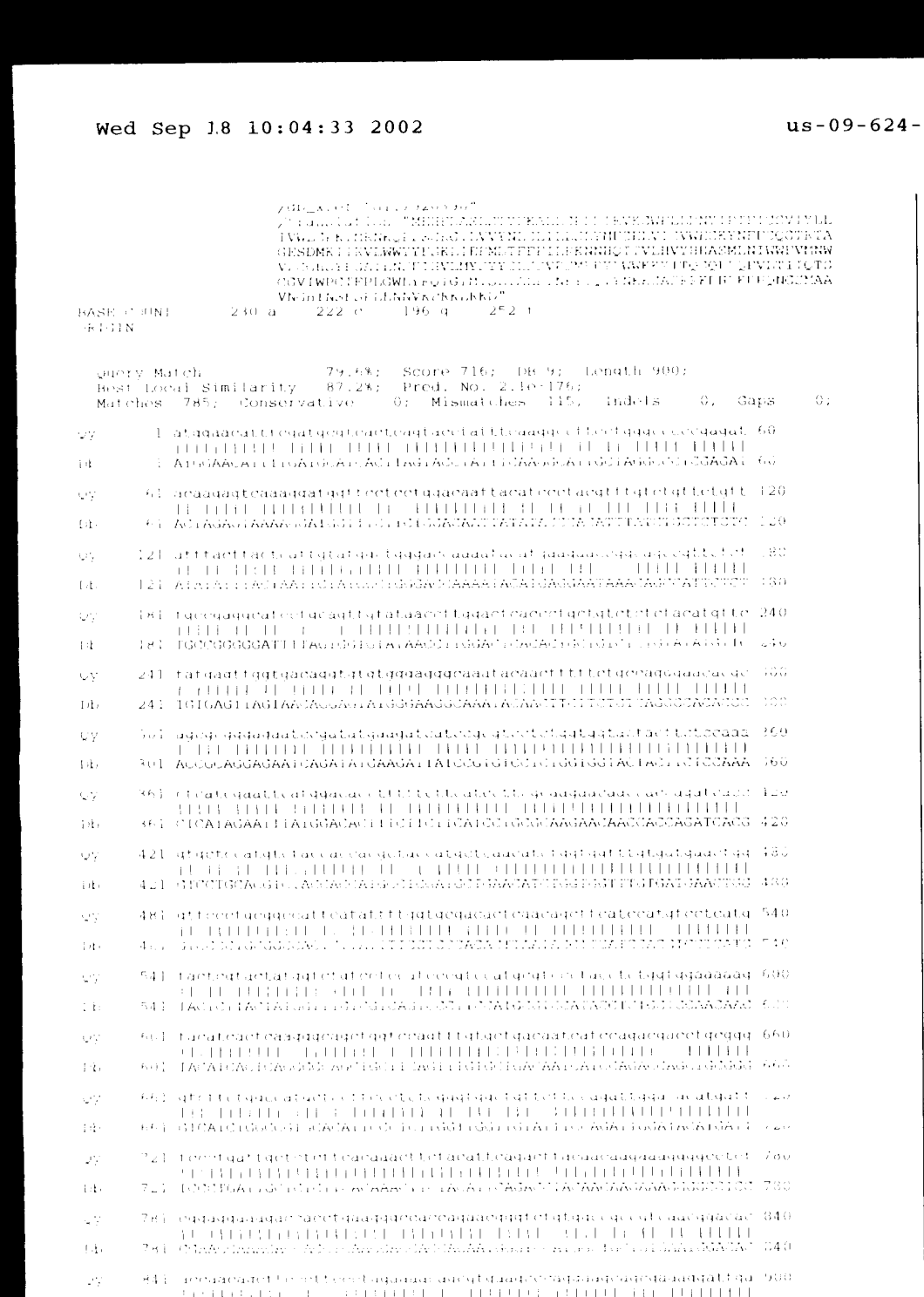


RESULT 4
AB071985
LOCUS AB071985 900 bp mRNA linear ROD 16-OCT-2001
DEFINITION Rattus norvegicus REL-C mRNA for fatty acid elongase 1, complete cds.
ACCESSION AB071985
VERSION AB071985.1 GI:16151798
KEYWORDS .
SOURCE Rattus norvegicus (strain:Sprague Dawley) male liver cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 900)
AUTHORS Inagaki,K., Aki,T., Fukuda,Y., Kawamoto,S., Shigeta,S., Ono,K. and







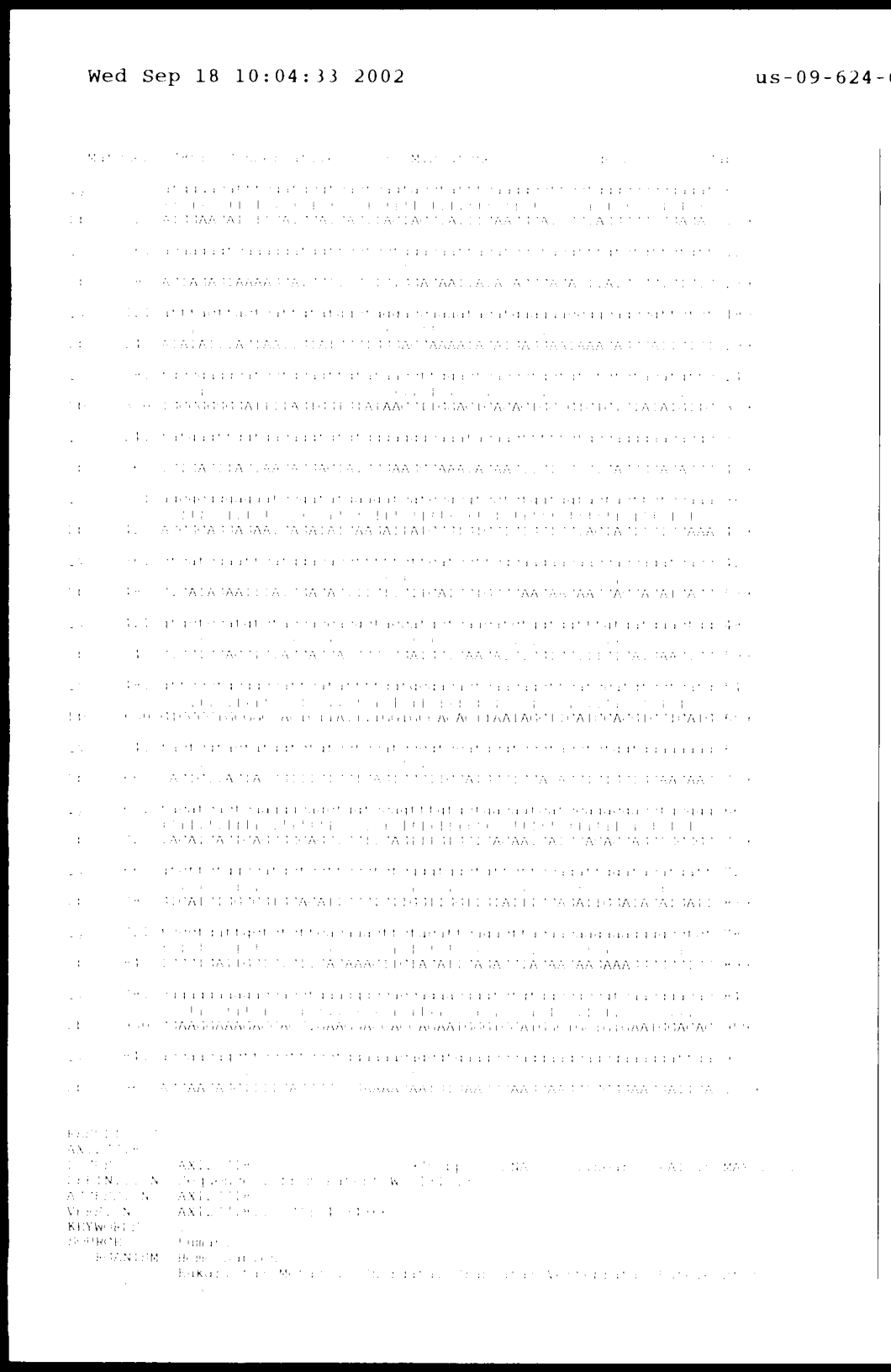


BASE COUNT 230 a 222 c 196 g 252 t
ORIGIN

Query Match 79.6%; Score 716; DB 9; Length 900;
Best Local Similarity 87.2%; Pred. No. 2.3e-176;
Matches 785; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

RESULT 6
HSM801903
LOCUS HSM801903 2757 bp mRNA linear PRI 10-MAR-2001
DEFINITION Homo sapiens cDNA, [NS: PRIp596B1824] (from clone DKFZp596B1824); complete cds.
ACCESSION AL136939
VERSION AL136939.1 GI:12053372
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2757)
AUTHORS Wiemann,S., Weil,B., Wellenreuther,R., Gassenhuber,J., Glassl,S., Ansorge,W., Boecher,M., Bloecker,H., Bauersachs,S., Blum,H., Lauber,J., Duesterhoeft,A., Beyer,A., Koehrer,K., Strack,N., Mewes,H.W., Ottenwaelder,B., Obermaier,B., Tampe,J., Heubner,D., Wambutt,R., Korn,B., Klein,M. and Poustka,A.
TITLE Toward a Catalog of Human Genes and Proteins: Sequencing and Analysis of 500 Novel Complete Protein Coding Human cDNAs
JOURNAL Genome Res. 11 (2), 422-435 (2001)
PUBMED 11230166
REFERENCE 2 (bases 1 to 2757)
AUTHORS Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (18 JAN 2000) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
Sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://www.rzpd.de and http://www.dkfz.de.
FEATURES Location/Qualifiers
source 1..2757
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="6p12.1-21.1"
/clone="DKFZp596B1824"
/tissue_type="uterus"
/clone_lib="596 (synonym: hutel). Vector pSport1; host DH10B; sites NotI + SalI/MluI"
/dev_stage="adult"
gene 120..1019
/gene="DKFZp596B1824"
CDS 120..1019
/gene="DKFZp596B1824"
/note="dJ483K16.1 extension of genmodel"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAB66373.1"
/db_xref="GI:12053373"
/translation="MAGLGFLIRVFGWFLLAGCLVFFCYVIFA
IVWLSHKYMKRKLPSCRSILVVVTGIDSGLGYMFCELVTCVWE
CRYNFTVCGLRLA
CSLDMKLRAVGAVIFSKLSPMCFFLTSFPPGLITVLEVYHHAC
MLNIWWFVMNW
VRGMSLLGAVLGLLAAVGWFSVMLFTLWWEKTITGSLLTPVF
VLLVTS
VNGHTNSTCPLENNVKPPKLRKD"
polyA_signal 2723..2728
polyA_site 2737
BASE COUNT 800 a 540 c 548 g 869 t
ORIGIN

Query Match 79.6%; Score 716; DB 9; Length 2757;
Best Local Similarity 87.2%; Pred. No. 2.6e-176;

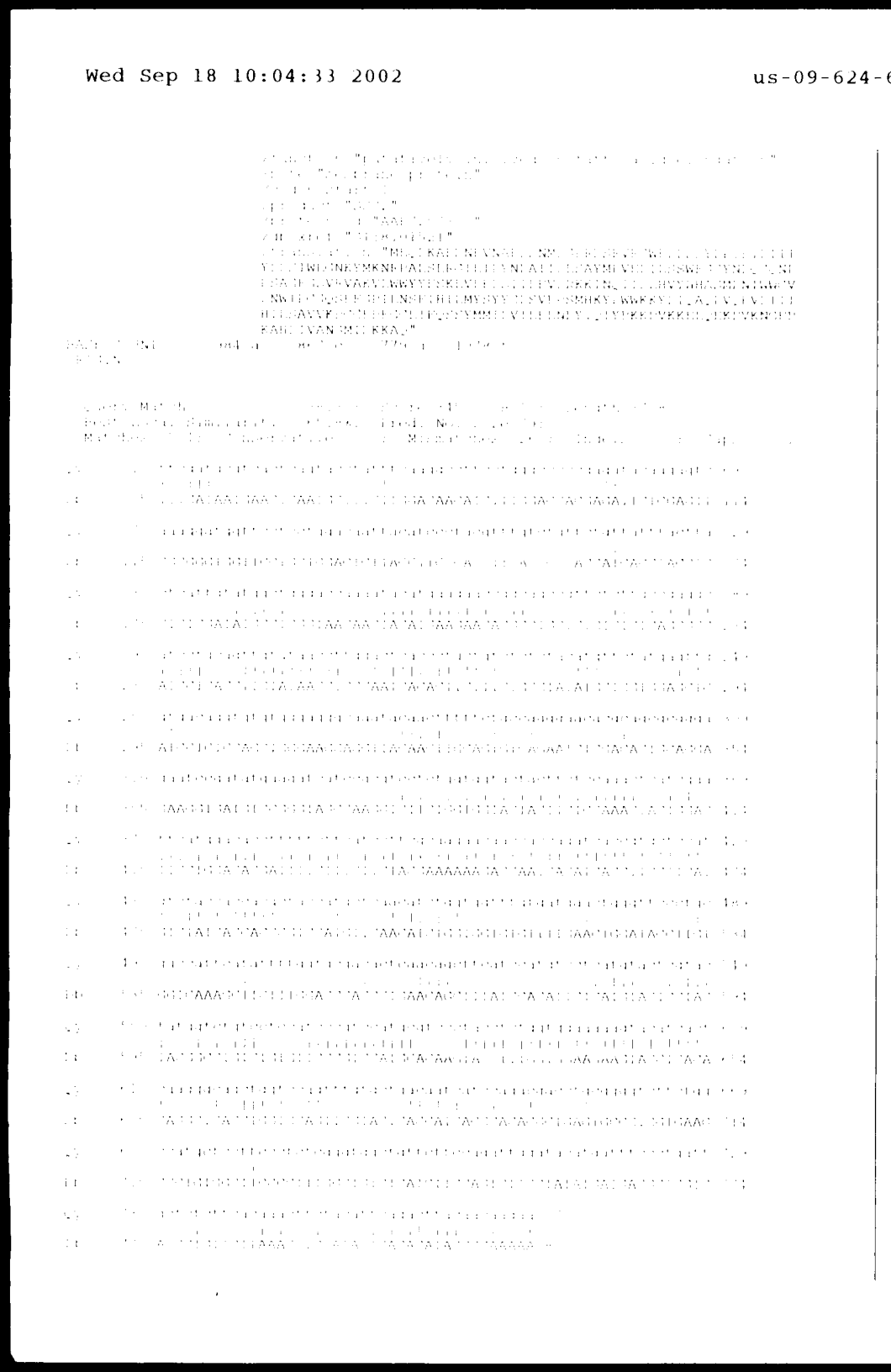


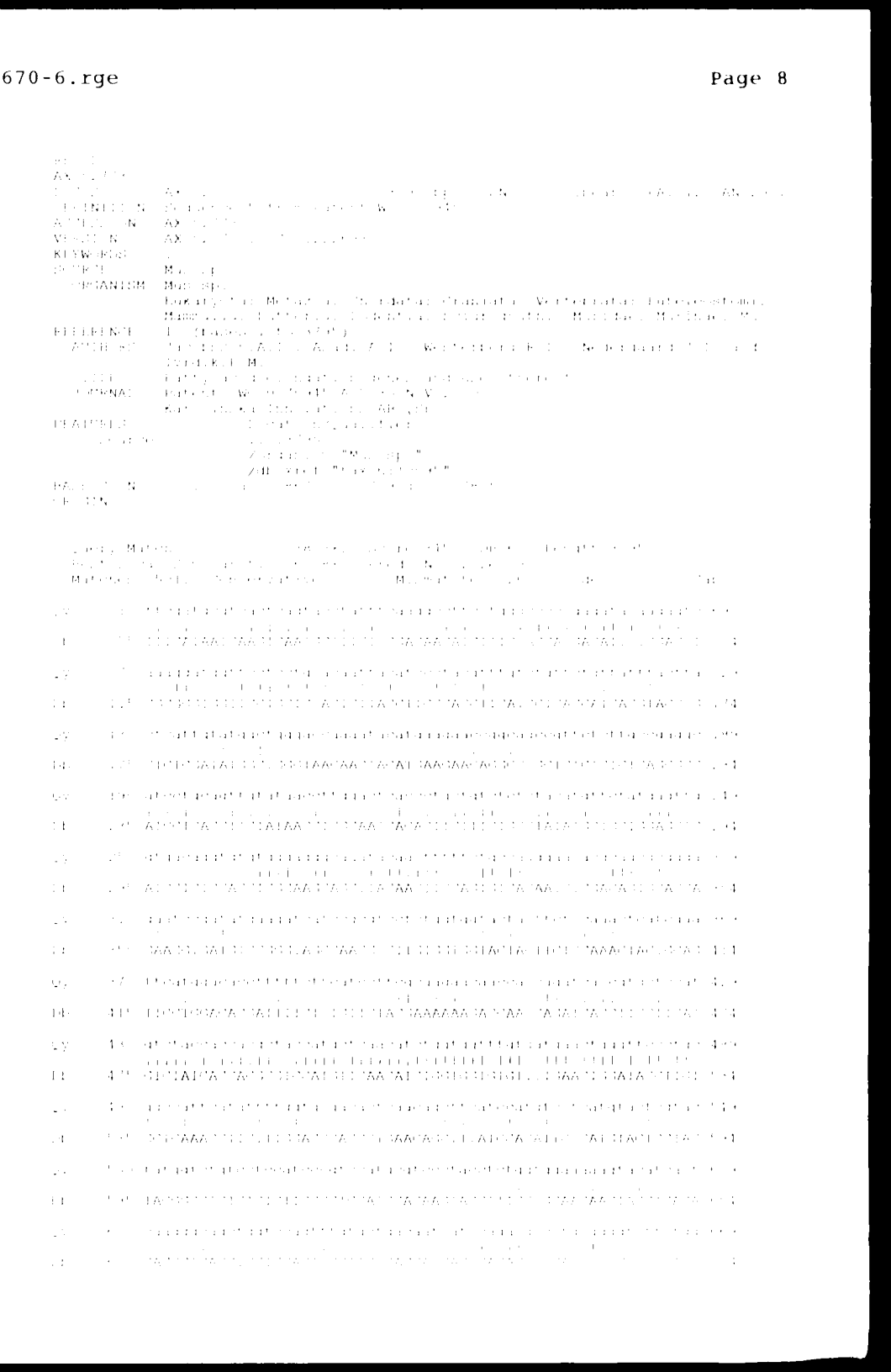
```
DEFINITION  Scophthalmus maximus polyunsaturated fatty acid elongase (elo)
            mRNA, complete cds.
ACCESSION   AF465624
VERSION     AF465624.1  GI:18448674
KEYWORDS    .
SOURCE      turbot.
  ORGANISM  Scophthalmus maximus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Pleuronectiformes;
            Pleuronectoidei; Scophthalmidae; Scophthalmus.
REFERENCE   1  (bases 1 to 1967)
  AUTHORS   Ayala,M.E., Fuentes,J.A. and Tocher,D.R.
  TITLE     Polyunsaturated fatty acid elongase from teleost fish
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1967)
  AUTHORS   Ayala,M.E., Fuentes,J.A. and Tocher,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JAN-2002) Institute of Aquaculture, University of
            Stirling, Stirling FK9 4LA, UK
FEATURES             Location/Qualifiers
     source          1..1967
                     /db_xref="taxon:52904"
                     /organism="Scophthalmus maximus"
                     /tissue_type="liver"
     gene            1..1967
                     /gene="Elo"
     5'UTR           1..28
                     /gene="Elo"
     CDS             29..913
                     /gene="Elo"
                     /codon_start=1
                     /product="polyunsaturated fatty acid elongase"
                     /protein_id="AAL69984.1"
                     /db_xref="GI:18448675"
                     /translation="MEFFHHFLNPYIDLVMKGKVRSLRVLMSDWPIFSTALTVWYLL
                     IVVMGFKVMNETNPHELSTSKKGLVVFNLGLTLFSFYMFYRLIVSAVSEGKVKGSVK
                     PEVDNKLNALSWYFPSLIFFSLGPGQITGLSVDYHTFHKAPIHKIMWVYDTC
                     IPCGHSYFIA"N"FVRVVA"""""F""A""TAIFT""AAPFY""LLLLSFLLHELIM
                     "AVGW""FFF""WCCC""CVVVLLLLLCNCYL""FRHNCLLPLRSLSAVPHCH
                     SNSIPSAHRMAHKKLRVP"
     3'UTR           914..1967
                     /gene="Elo"
BASE COUNT      440 a      554 c      592 g      381 t
ORIGIN
```

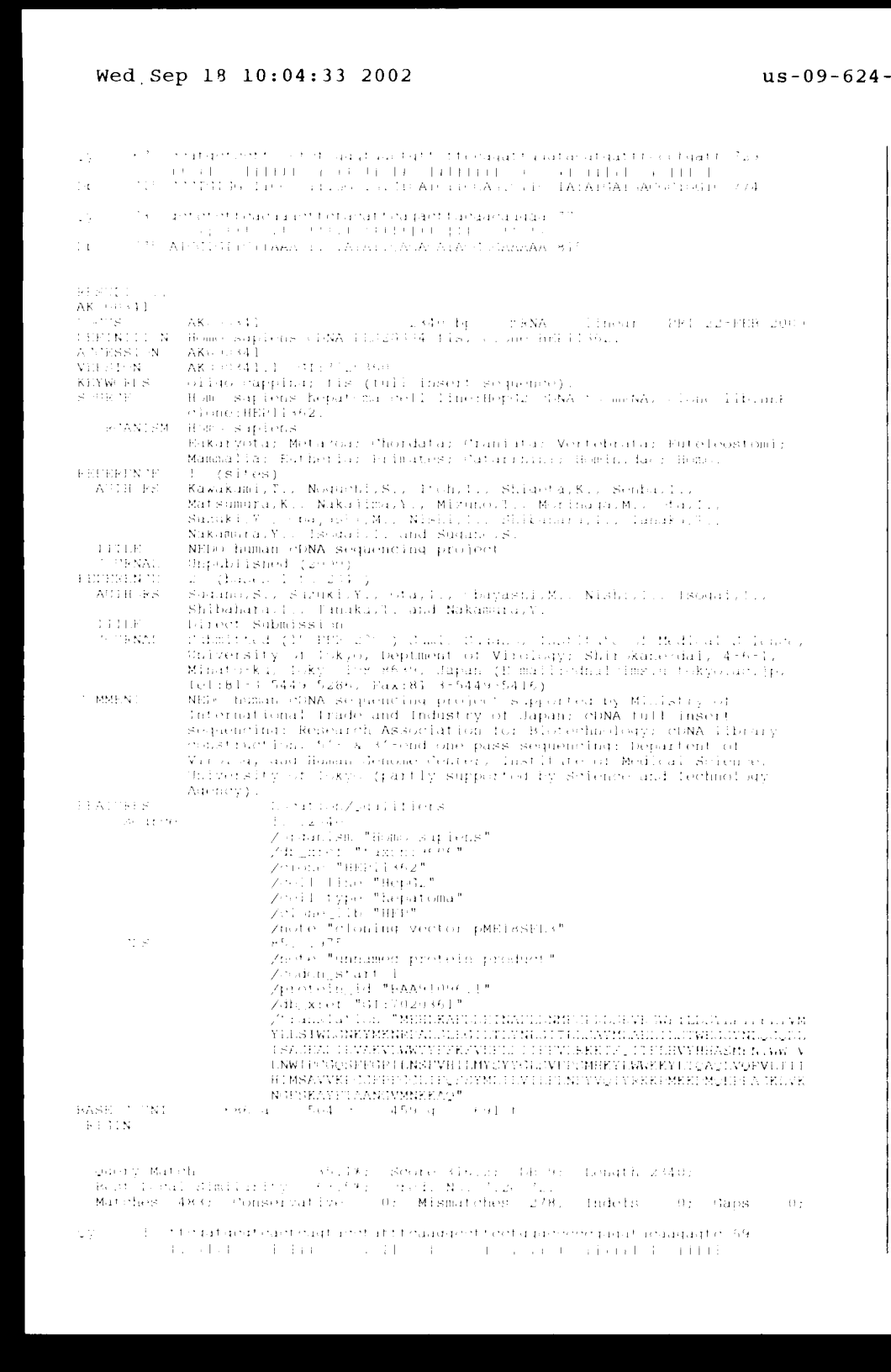
```
  1 atgtttcatcctttcctcaatccctacatcgacctcgtcatgaaaggcaaggtccgctcgc
 61 tgcgtgctcatgtccgactggcccatcttctccacggccctcacggtctggtacctcctc
121 atcgtcgtcatgggcttcaaggtcatgaacgagaccaacccccacgagctgtccacgtcc
181 aagaagggcctcgtcgtcttcaacctcggcctcaccctcttctccttctacatgttctac
241 cgcctcatcgtctccgccgtctccgagggcaaggtcaagggctccgtcaagcccgaggtc
301 gacaacaagctcaacgccctctcctggtacttccccagcctcatcttcagcctcggccct
361 ggccagatcaccggcctctccgtcgactaccacacgttccacaaggcccccatccacaag
421 atcatgtgggtctacgacacctgcatcccctgcggccactcctacttcatcgccaacttc
481 gtccgcgtcgtcgccgccttcttcctcttcgccgccatcttcaccgccgccccgttctac
541 ctcctctccgccgtcgccgagggcaaggtcaagggctccgtcaagcccgaggtcgacaac
601 aagctcaacgccctctcctggtacttccccagcctcatcttcagcctcggccctggccag
661 atcaccggcctctccgtcgactaccacacgttccacaaggcccccatccacaagatcatg
721 tgggtctacgacacctgcatcccctgcggccactcctacttcatcgccaacttcgtccgc
781 gtcgtcgccgccttcttcctcttcgccgccatcttcaccgccgccccgttctacctcctc
841 tccgccgtcgccgagggcaaggtcaagggctccgtcaagcccgaggtcgacaacaagctc
901 aacgccctctcctggtacttccccagcctcatcttcagcctcggccctggccagatcacc
961 ggcctctccgtcgactaccacacgttccacaaggcccccatccacaagatcatgtgggtc
1021 tacgacacctgcatcccctgcggccactcctacttcatcgccaacttcgtccgcgtcgtc
1081 gccgccttcttcctcttcgccgccatcttcaccgccgccccgttctacctcctctccgcc
1141 gtcgccgagggcaaggtcaagggctccgtcaagcccgaggtcgacaacaagctcaacgcc
1201 ctctcctggtacttccccagcctcatcttcagcctcggccctggccagatcaccggcctc
1261 tccgtcgactaccacacgttccacaaggcccccatccacaagatcatgtgggtctacgac
1321 acctgcatcccctgcggccactcctacttcatcgccaacttcgtccgcgtcgtcgccgcc
1381 ttcttcctcttcgccgccatcttcaccgccgccccgttctacctcctctccgccgtcgcc
1441 gagggcaaggtcaagggctccgtcaagcccgaggtcgacaacaagctcaacgccctctcc
1501 tggtacttccccagcctcatcttcagcctcggccctggccagatcaccggcctctccgtc
1561 gactaccacacgttccacaaggcccccatccacaagatcatgtgggtctacgacacctgc
1621 atcccctgcggccactcctacttcatcgccaacttcgtccgcgtcgtcgccgccttcttc
1681 ctcttcgccgccatcttcaccgccgccccgttctacctcctctccgccgtcgccgagggc
1741 aaggtcaagggctccgtcaagcccgaggtcgacaacaagctcaacgccctctcctggtac
1801 ttccccagcctcatcttcagcctcggccctggccagatcaccggcctctccgtcgactac
1861 cacacgttccacaaggcccccatccacaagatcatgtgggtctacgacacctgcatcccc
1921 tgcggccactcctacttcatcgccaacttcgtccgcgtcgtcgccgccttcttcctcttc
1981 gccgccatcttcaccgccgccccgttctacctcctctccgccgtcgccgagggcaaggtc
2041 aagggctccgtcaagcccgaggtcgacaacaagctcaacgccctctcctggtacttcccc
2101 agcctcatcttcagcctcggccctggccagatcaccggcctctccgtcgactaccacacg
2161 ttccacaaggcccccatccacaagatcatgtgggtctacgacacctgcatcccctgcggc
2221 cactcctacttcatcgccaacttcgtccgcgtcgtcgccgccttcttcctcttcgccgcc
2281 atcttcaccgccgccccgttctacctcctctccgccgtcgccgagggcaaggtcaagggc
2341 tccgtcaagcccgaggtcgacaacaagctcaacgccctctcctggtacttccccagcctc
2401 atcttcagcctcggccctggccagatcaccggcctctccgtcgactaccacacgttccac
2461 aaggcccccatccacaagatcatgtgggtctacgacacctgcatcccctgcggccactcc
2521 tacttcatcgccaacttcgtccgcgtcgtcgccgccttcttcctcttcgccgccatcttc
2581 accgccgccccgttctacctcctctccgccgtcgccgagggcaaggtcaagggctccgtc
2641 aagcccgaggtcgacaacaagctcaacgccctctcctggtacttccccagcctcatcttc
2701 agcctcggccctggccagatcaccggcctctccgtcgactaccacacgttccacaaggcc
2761 cccatccacaagatcatgtgggtctacgacacctgcatcccctgcggccactcctacttc
2821 atcgccaacttcgtccgcgtcgtcgccgccttcttcctcttcgccgccatcttcaccgcc
2881 gccccgttctacctcctctccgccgtcgccgagggcaaggtcaagggctccgtcaagccc
2941 gaggtcgacaacaagctcaacgccctctcctggtacttccccagcctcatcttcagcctc
3001 ggccctggccagatcaccggcctctccgtcgactaccacacgttccacaaggcccccatc
3061 cacaagatcatgtgggtctacgacacctgcatcccctgcggccactcctacttcatcgcc
3121 aacttcgtccgcgtcgtcgccgccttcttcctcttcgccgccatcttcaccgccgccccg
3181 ttctacctcctctccgccgtcgccgagggcaaggtcaagggctccgtcaagcccgaggtc
3241 gacaacaagctcaacgccctctcctggtacttccccagcctcatcttcagcctcggccct
3301 ggccagatcaccggcctctccgtcgactaccacacgttccacaaggcccccatccacaag
3361 atcatgtgggtctacgacacctgcatcccctgcggccactcctacttcatcgccaacttc
3421 gtccgcgtcgtcgccgccttcttcctcttcgccgccatcttcaccgccgccccgttctac
3481 ctcctctccgccgtcgccgagggcaaggtcaagggctccgtcaagcccgaggtcgacaac
3541 aagctcaacgccctctcctggtacttccccagcctcatcttcagcctcggccctggccag
3601 atcaccggcctctccgtcgactaccacacgttccacaaggcccccatccacaagatcatg
3661 tgggtctacgacacctgcatcccctgcggccactcctacttcatcgccaacttcgtccgc
3721 gtcgtcgccgccttcttcctcttcgccgccatcttcaccgccgccccgttctacctcctc
3781 tccgccgtcgccgagggcaaggtcaagggctccgtcaagcccgaggtcgacaacaagctc
3841 aacgccctctcctggtacttccccagcctcatcttcagcctcggccctggccagatcacc
3901 ggcctctccgtcgactaccacacgttccacaaggcccccatccacaagatcatgtgggtc
3961 tacgac
```

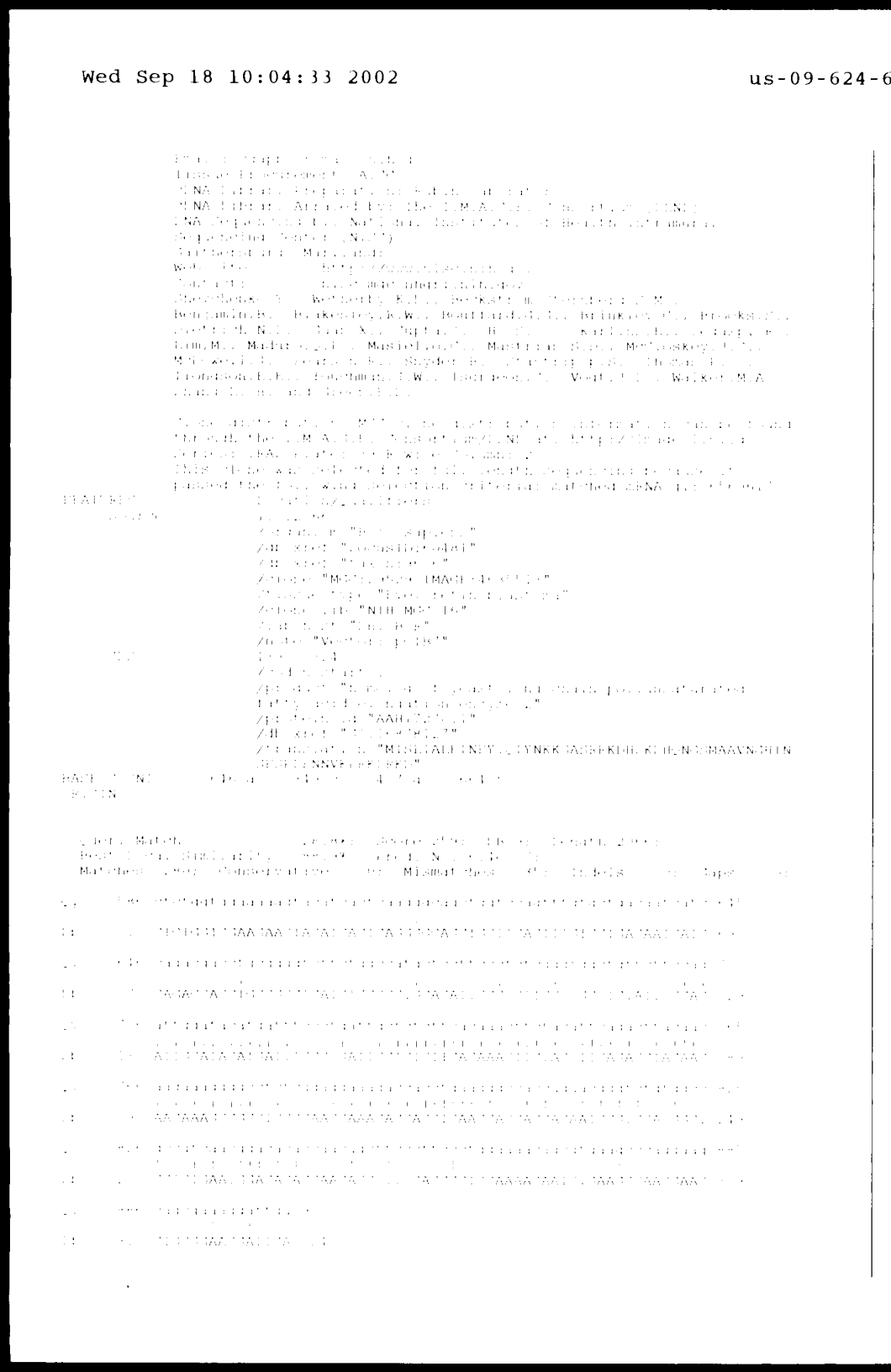
```
LOCUS       AF465624                1967 bp    mRNA    linear   ROD 06-MAY-2002
DEFINITION  Scophthalmus maximus polyunsaturated fatty acid elongase (elo)
            mRNA, complete cds.
ACCESSION   AF465624
VERSION     AF465624.1  GI:18448674
KEYWORDS    .
SOURCE      turbot.
  ORGANISM  Scophthalmus maximus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Pleuronectiformes;
            Pleuronectoidei; Scophthalmidae; Scophthalmus.
REFERENCE   1  (bases 1 to 1967)
  AUTHORS   Ayala,M.E., Fuentes,J.A. and Tocher,D.R.
  TITLE     Polyunsaturated fatty acid elongase from teleost fish
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1967)
  AUTHORS   Ayala,M.E., Fuentes,J.A. and Tocher,D.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JAN-2002) Institute of Aquaculture, University of
            Stirling, Stirling FK9 4LA, UK
FEATURES             Location/Qualifiers
     source          1..1967
                     /db_xref="taxon:52904"
                     /organism="Scophthalmus maximus"
                     /tissue_type="liver"
     gene            1..1967
                     /gene="Elo"
     5'UTR           1..28
                     /gene="Elo"
     CDS             29..913
                     /gene="Elo"
                     /codon_start=1
                     /product="polyunsaturated fatty acid elongase"
                     /protein_id="AAL69984.1"
                     /db_xref="GI:18448675"
                     /translation="MEFFHHFLNPYIDLVMKGKVRSLRVLMSDWPIFSTALTVWYLL
                     IVVMGFKVMNETNPHELSTSKKGLVVFNLGLTLFSFYMFYRLIVSAVSEGKVKGSVK
                     PEVDNKLNALSWYFPSLIFFSLGPGQITGLSVDYHTFHKAPIHKIMWVYDTC
                     IPCGHSYFIA"N"FVRVVA"""""F""A""TAIFT""AAPFY""LLLLSFLLHELIM
                     "AVGW""FFF""WCCC""CVVVLLLLLCNCYL""FRHNCLLPLRSLSAVPHCH
                     SNSIPSAHRMAHKKLRVP"
     3'UTR           914..1967
                     /gene="Elo"
BASE COUNT      440 a      554 c      592 g      381 t
ORIGIN
```

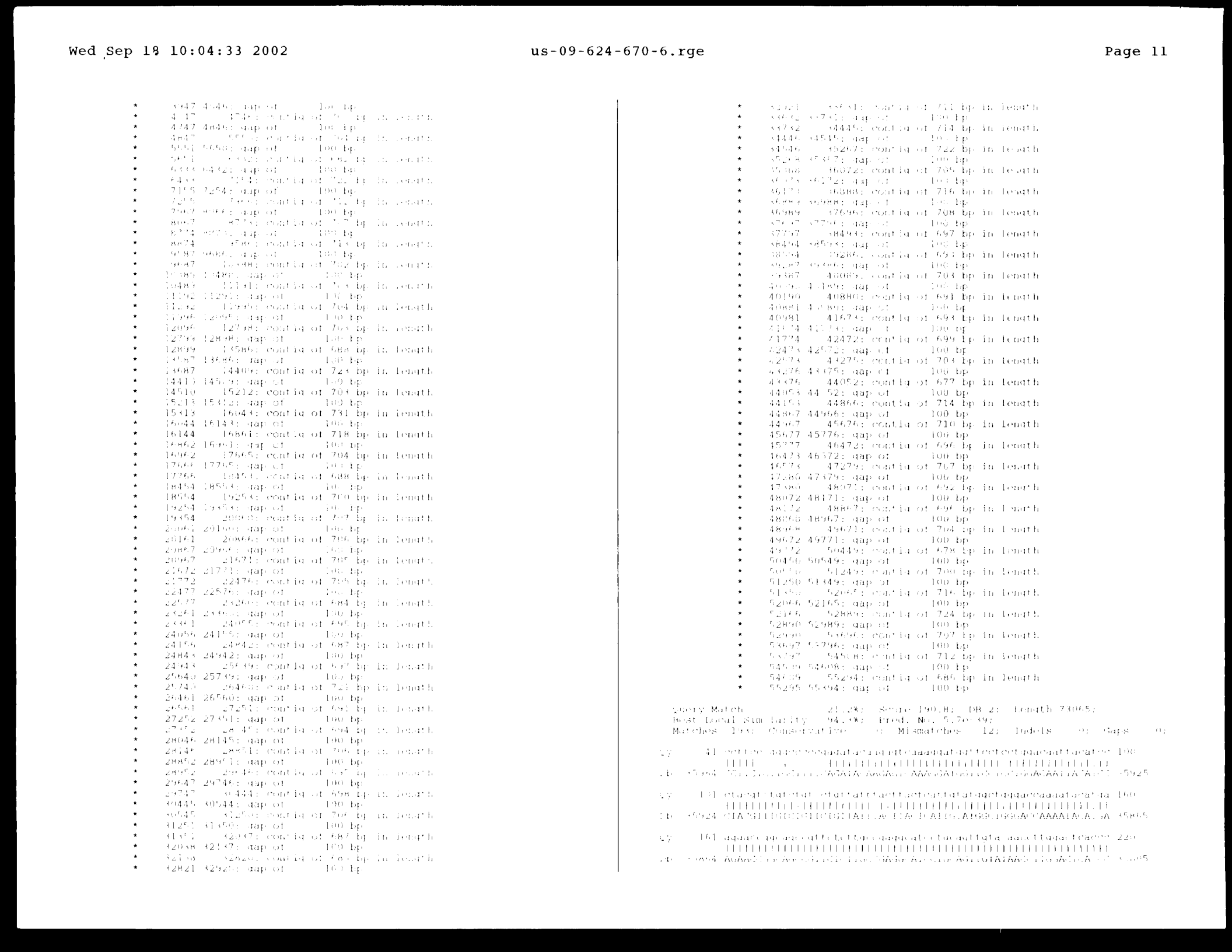
```
LOCUS       AF176648                3708 bp    mRNA    linear   ROD 06-MAY-2000
DEFINITION  Mus musculus Elovl2 mRNA, complete cds.
ACCESSION   AF176648
VERSION     AF176648.1  GI:8670528
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3708)
  AUTHORS   Tvrdik,P., Westerberg,R., Silve,S., Asadi,A., Jakobsson,A.,
            Cannon,B., Loison,G. and Jacobsson,A.
  TITLE     Role of a new mammalian gene family in the biosynthesis of very
            long chain fatty acids and sphingolipids
  JOURNAL   J. Cell Biol. 149 (4), 707-718 (2000)
  MEDLINE   20256178
  PUBMED    10791985
REFERENCE   2  (bases 1 to 3708)
  AUTHORS   Tvrdik,P., Westerberg,R., Silve,S., Asadi,A., Jakobsson,A.,
            Cannon,B., Loison,G. and Jacobsson,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUL-1999) The Arrhenius Laboratories F3, The
            Wenner-Gren Institute, Stockholm University, Stockholm S-106 91,
            Sweden
FEATURES             Location/Qualifiers
     source          1..3708
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /sex="male"
                     /tissue_type="liver"
                     /dev_stage="6-12 weeks"
     gene            1..3708
                     /gene="Ssc2"
     CDS             57..915
                     /gene="Ssc2"
```

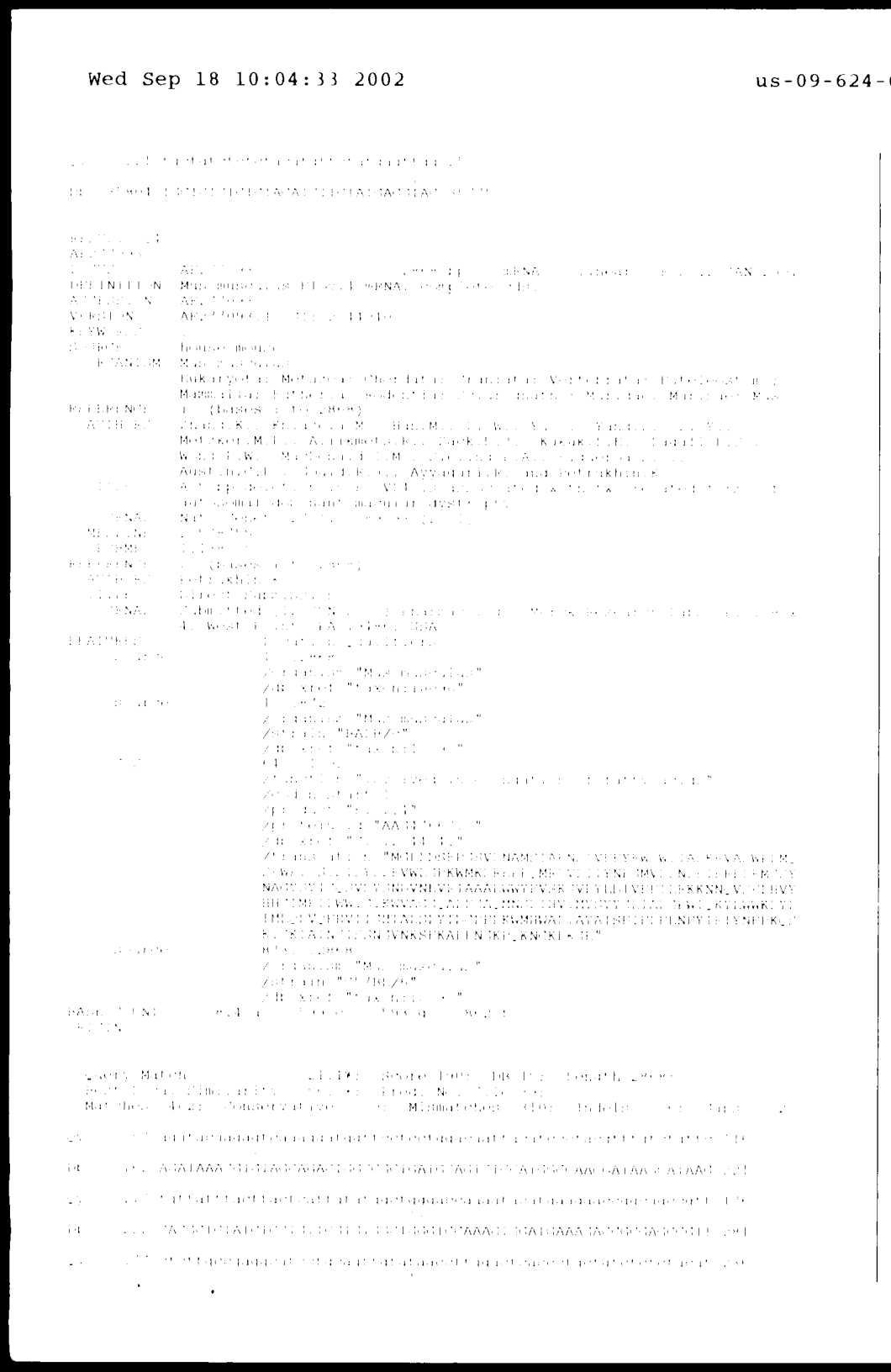












```
                  Institute (BRI) (supported by Japan
                  Technology Center of Co); 5' & 3' end sequencing: RAB,
                  and Biotechnology Center, National Institute of Technology and
                  Evaluation; clone selection for full insert sequencing: RAB and
                  BRI.
FEATURES             Location/Qualifiers
     source          1..2219
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RPCI98-05e6"
                     /tissue_type="brain"
                     /clone_lib="RPCI-"
                     /dev_stage="fetus"
                     /note="Vector: pMC1853."
     CDS             992..1240
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="BAB70805.1"
                     /db_xref="GI:16549978"
                     /translation="MGLLISSDVSGVLNVVSIALNEIYGFYEKEVEMGLLMQ
                     SIWFTLSISILYLLFVWLCFKWMKPRRPVPMPEVLVIYSPCMVLNLFTPHSLPMGSY
                     NAGYGFTGQWVFLNNVGLVKIAAALWFTVSKSVGLLVFLVRRKANQVSFHHV
                     HHFIMFFLWGCSAEWVSGLACFGLNSFVHVIMYSYYGLSAIGPSIQKYLWWGRYL
                     TRFQFVQFHVTIGHTALSLYTPDFPWMNWAIGAYAICFLFLFNEFYRKYLKRKP
                     KAEKIAMNIGLGNIVFPFLIMISCKEVRYGNCKAGD"
BASE COUNT       504 a    451 c    486 g    679 t
ORIGIN

  Query Match             14.8%;  Score 174;  DB 9;  Length 2219;
  Best Local Similarity   54.5%;  Pred. No. 1.2e-54;
  Matches  302;  Conservative    0;  Mismatches  252;  Indels   66;  Gaps   27;

Qy      1 ATATAAAGGTAGGTAAAATTATTGTTTCTGTGAATTATTCATCTTTTGTTAATTGTATTGC 116
          ||||  | ||  ||   | |  |  | |  |      ||||  |
Db      4 AGAGAAGCGTGTGCAAAAGTGGCATGTGATGTATCTGTGCACATAAGTAATAAAGC 479

Qy    117 TTTATTGACTTACTGTATTATGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 176
          | ||||  | ||  |||| ||||||| ||||||  ||||||  ||||  | || ||
Db    480 TATGTTGATTTCTGTGTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 539

Qy    177 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAT 236
                 ||     | ||     | ||||  ||||  |     ||||    ||||  ||
Db    540 TTTGATGTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAT 599

Qy    237 TTTTATGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 296
          | |  || |          |   |  | |    ||| | | | | ||||| | |
Db    600 TTTCAGAGAGTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 659

Qy    297 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 356
                  |          |   |  |  |  |  |  |  ||  | ||||||  |
Db    660 AGATTATTTCTAATAATGTTCATGAAGTTCAGCATAAGTTTTTTTTTTTTTTTTTTTTT 719

Qy    357 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 416
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db    720 AGTTTCTTTCCTTCATGTGTATCACTCTGTACGATGTTTACCTTGTGGTGCATTGGAAT 779

Qy    417 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 476
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db    780 TAAGTGTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 839

Qy    477 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 536
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db    840 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 899

Qy    537 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 596
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db    900 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 959

Qy    597 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 656
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db    960 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1019

Qy    657 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 716
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1020 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1079

Qy    717 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 776
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1080 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1139

Qy    777 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 836
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1140 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1199

Qy    837 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 896
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1200 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1259

Qy    897 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 956
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1260 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1319

Qy    957 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1016
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1320 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1379

Qy   1017 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1076
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1380 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1439

Qy   1077 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1136
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1440 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1499

Qy   1137 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1196
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1500 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1559

Qy   1197 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1256
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1560 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1619

Qy   1257 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1316
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1620 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1679

Qy   1317 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1376
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1680 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1739

Qy   1377 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1436
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1740 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1799

Qy   1437 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1496
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1800 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1859

Qy   1497 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1556
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1860 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1919

Qy   1557 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1616
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1920 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1979

Qy   1617 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1676
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   1980 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2039

Qy   1677 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1736
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   2040 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2099

Qy   1737 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1796
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   2100 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2159

Qy   1797 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1856
          |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db   2160 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2219

Search completed: September 17, 2002, 18:44:18
Job time : 4295.608
```

Geneseq version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

GM protein - protein search, using sw model

Run on: September 17, 2002, 18:47:26 ; Search time 29.2 Seconds
(without alignments)
5424.687 Million cell updates/sec

Title: US-09-624-670-6
Perfect score: 900
Sequence: 1 atgaaacattcaatgcgtc.....aaaacaacgaaaaaaattaa 900

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 958457221 residues

Total number of hits satisfying chosen parameters: 347287

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1985.DAT:*
2: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1986.DAT:*
3: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1987.DAT:*
4: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1988.DAT:*
5: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1989.DAT:*
6: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1990.DAT:*
7: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1991.DAT:*
8: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1992.DAT:*
9: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1993.DAT:*
10: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1994.DAT:*
11: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1995.DAT:*
12: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1996.DAT:*
13: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1997.DAT:*
14: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1998.DAT:*
15: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA1999.DAT:*
16: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA2000.DAT:*
17: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA2001.DAT:*
18: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA2002.DAT:*
19: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA2003.DAT:*
20: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA2004.DAT:*
21: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA2005.DAT:*
22: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA2006.DAT:*
23: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA2007.DAT:*
24: /cgn2_6/ptodata/1/na/N_Geneseq_032802/NA2008.DAT:*

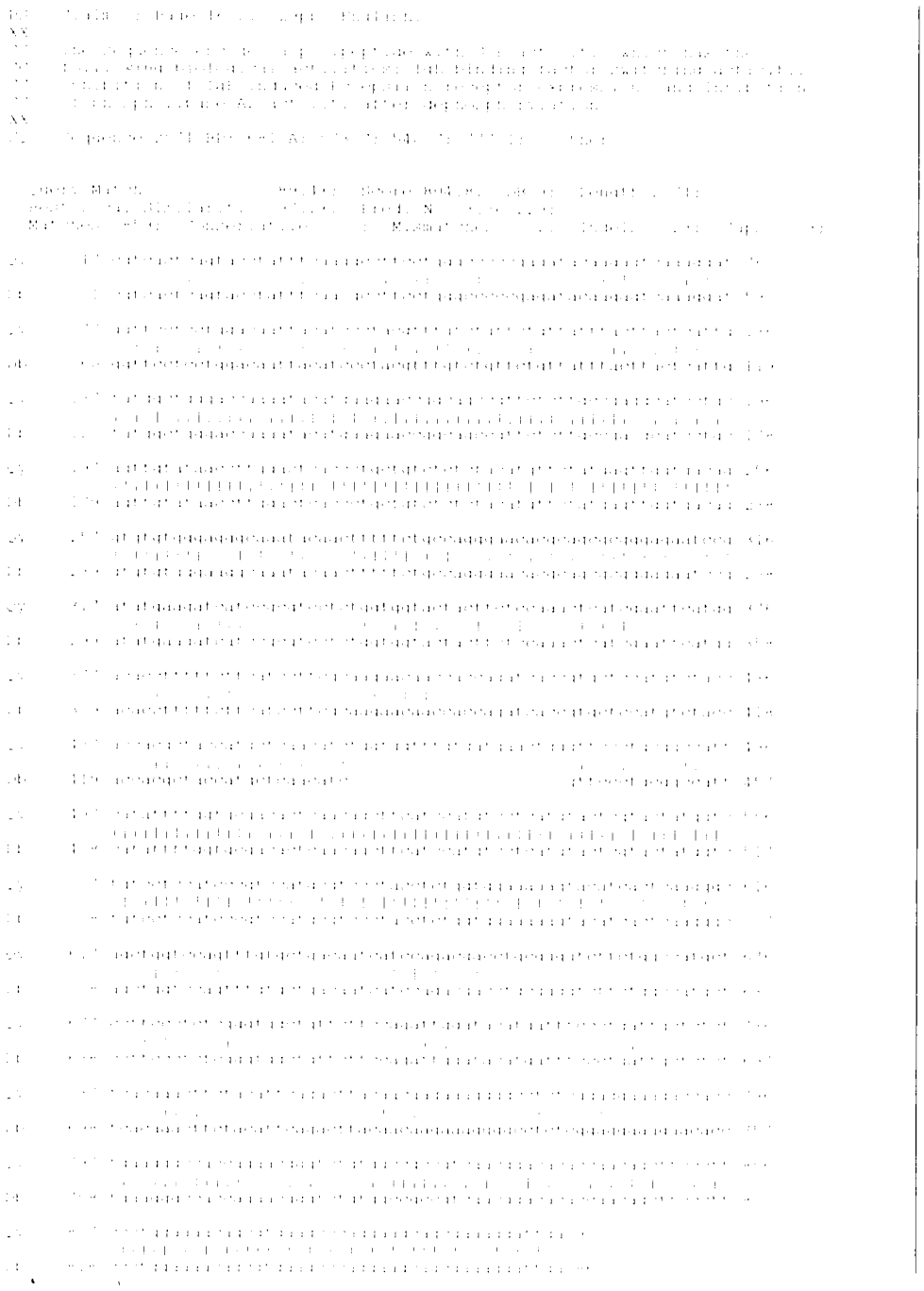
Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

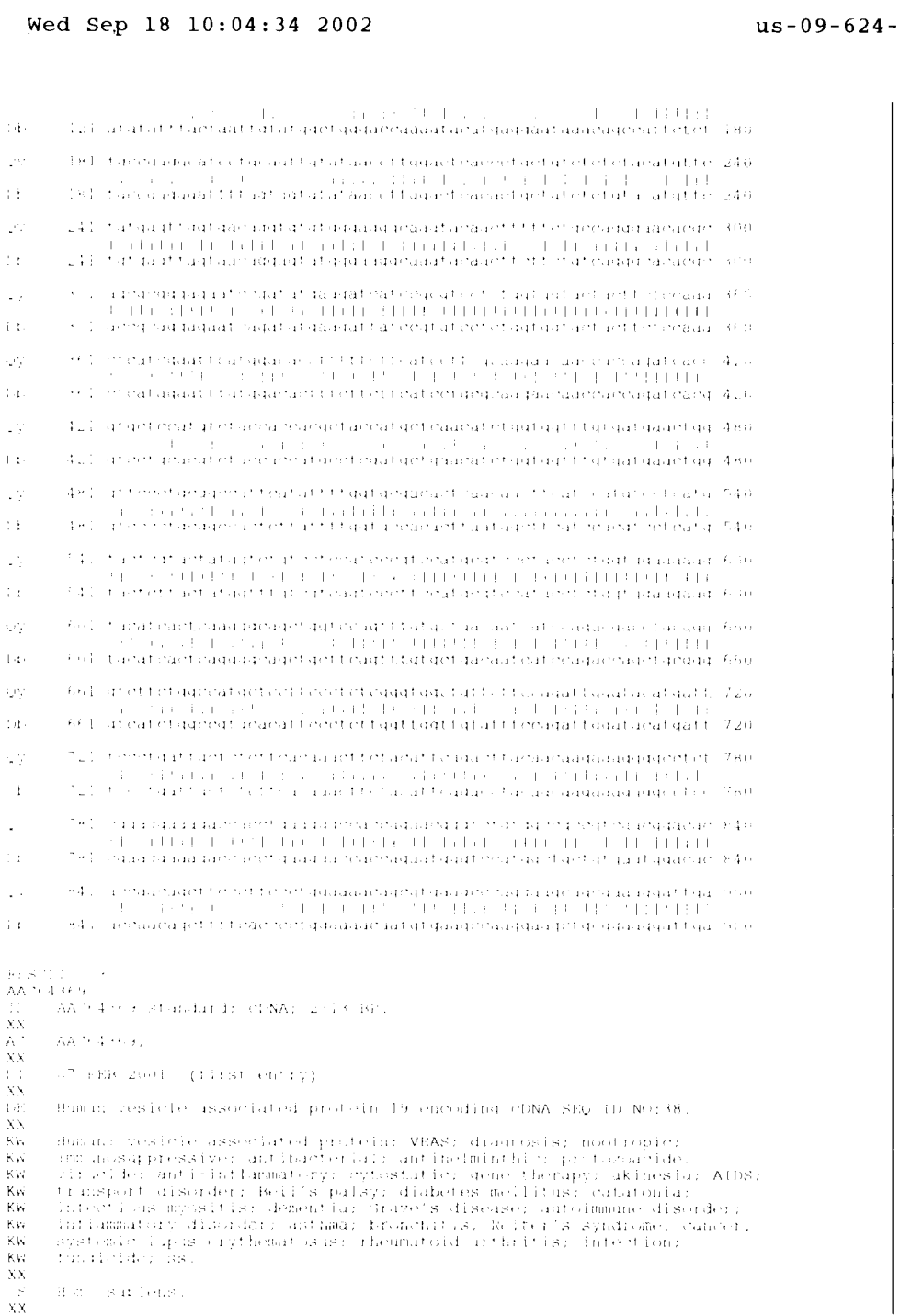
SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	804.8	89.4	2571	9	AAN81168	Fragment from clon
2	716	79.6	914	21	AAZ94241	Human clonqaso HSE
3	716	79.6	2313	21	AAC94359	Human vesicle-asso
4	716	79.6	2627	22	AAK94320	Human full-length
5	716	79.6	2769	22	AAS44747	Human full-length
6	716	79.6	3083	22	AAS44919	Human contig polyn
7	714.4	79.4	2713	21	AAA54182	Breast cancer prot
8	541.6	59.4	675	22	AAD07290	Human cold-induced
9	519.4	57.7	871	22	AAS27318	cDNA encoding nove
10	475	52.8	680	22	AAK91944	Human cDNA 5' end
11	475	52.8	680	22	AAK93949	Human cDNA clone r
12	465.6	51.7	748	22	AAS27317	cDNA encoding nove
13	451.2	50.1	817	22	AAK94092	Human cDNA clone r
14	368.6	41.0	809	22	AAI97437	Human neuroblastom
15	364	40.4	795	22	AAI95727	Human neuroblastom
16	361.6	40.2	927	22	AAK93155	Human cDNA clone r
17	345	38.3	3795	22	AAC91702	Mouse fatty acid e
18	340.2	37.8	2225	21	AAC77854	Human cancer assoc
19	299.2	33.2	874	22	AAK91805	Human cDNA 5'-end
20	299.2	33.2	874	22	AAK93947	Human cDNA clone r
21	299.2	33.2	2290	22	AAK94230	Human full-length
22	265	29.4	488	23	AAS76145	DNA encoding novel
23	174	19.3	1203	23	AAS91595	DNA encoding novel
c 24	145.2	16.1	196	22	ABA48732	Human breast cell
c 25	145.2	16.1	196	22	ABA66640	Human foetal liver
c 26	145.2	16.1	196	22	ABA33706	Probe #12172 for q
c 27	145.2	16.1	196	22	AAK15070	Human brain expres
c 28	145.2	16.1	196	22	AAK40797	Human bone marrow
c 29	145.2	16.1	196	22	AAI21567	Probe #11500 for q
c 30	145.2	16.1	196	22	AAI46858	Probe #15544 used
c 31	145.2	16.1	196	22	AAI07261	Probe #7252 used t
c 32	145.2	16.1	212	22	ABA49657	Human breast cell
c 33	145.2	16.1	212	22	ABA67559	Human foetal liver
c 34	145.2	16.1	212	22	AAK15982	Human brain expres
c 35	145.2	16.1	212	22	AAK41727	Human bone marrow
c 36	145.2	16.1	212	22	AAI22484	Probe #12417 for q
c 37	145.2	16.1	212	22	AAI47775	Probe #16461 used
c 38	145.2	16.1	212	22	AAI08169	Probe #8160 used t
c 39	145.2	16.1	409	22	ABA51459	Human breast cell
c 40	145.2	16.1	409	22	AAI24350	Probe #14283 for q
c 41	145.2	16.1	460	22	ABA49624	Human breast cell
c 42	145.2	16.1	480	22	ABA67624	Human foetal liver
c 43	145.2	16.1	480	22	ABA33624	Probe #2289 for qe
c 44	145.2	16.1	480	22	AAK94240	Human brain expres
c 45	145.2	16.1	480	22	AAK47787	Human bone marrow

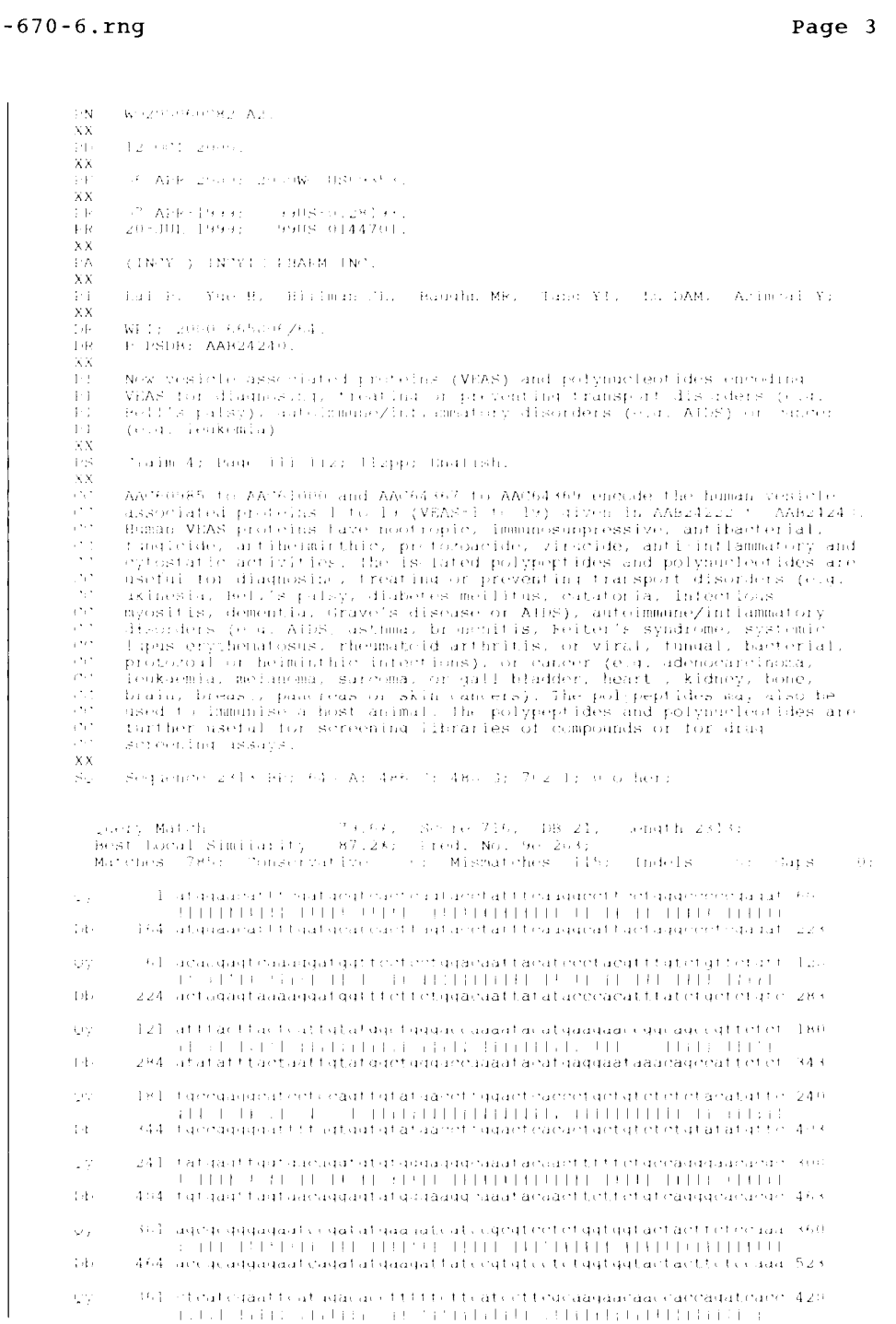
ALIGNMENTS

RESULT 1
AAN81168
ID AAN81168 standard; DNA; 2571 BP.
XX
AC AAN81168;
XX
DT 06-NOV-1900 (first entry)
XX
DE Fragment from clone encoding polypeptide with glycosylation inhibiting
DE factor activity.
XX
KW Glycosylation; GIF; IGIP4; IgE; anti-inflammatory agent; ss.
XX
PN EP285405 A.
XX
PD 05-OCT-1988.
XX
PF 06-MAR-1908; 88EP-0302852.
XX
PR 31-MAR-1987; 87US-0032859.
XX
PA (UYJO) JOHN HOPKINS UNIVERSITY.
XX
PI Ishizaka K, Marone GL, Meno KW;
XX
DR WPI; 1988-301145/40.
DR P-PSDB; AAP82590.
XX
PI Nucleic acid encoding glycosylation inhibiting peptide - capable of
PI suppressing IgE immunoglobulin responses, for use as antiinflammatory
PI agents.
XX

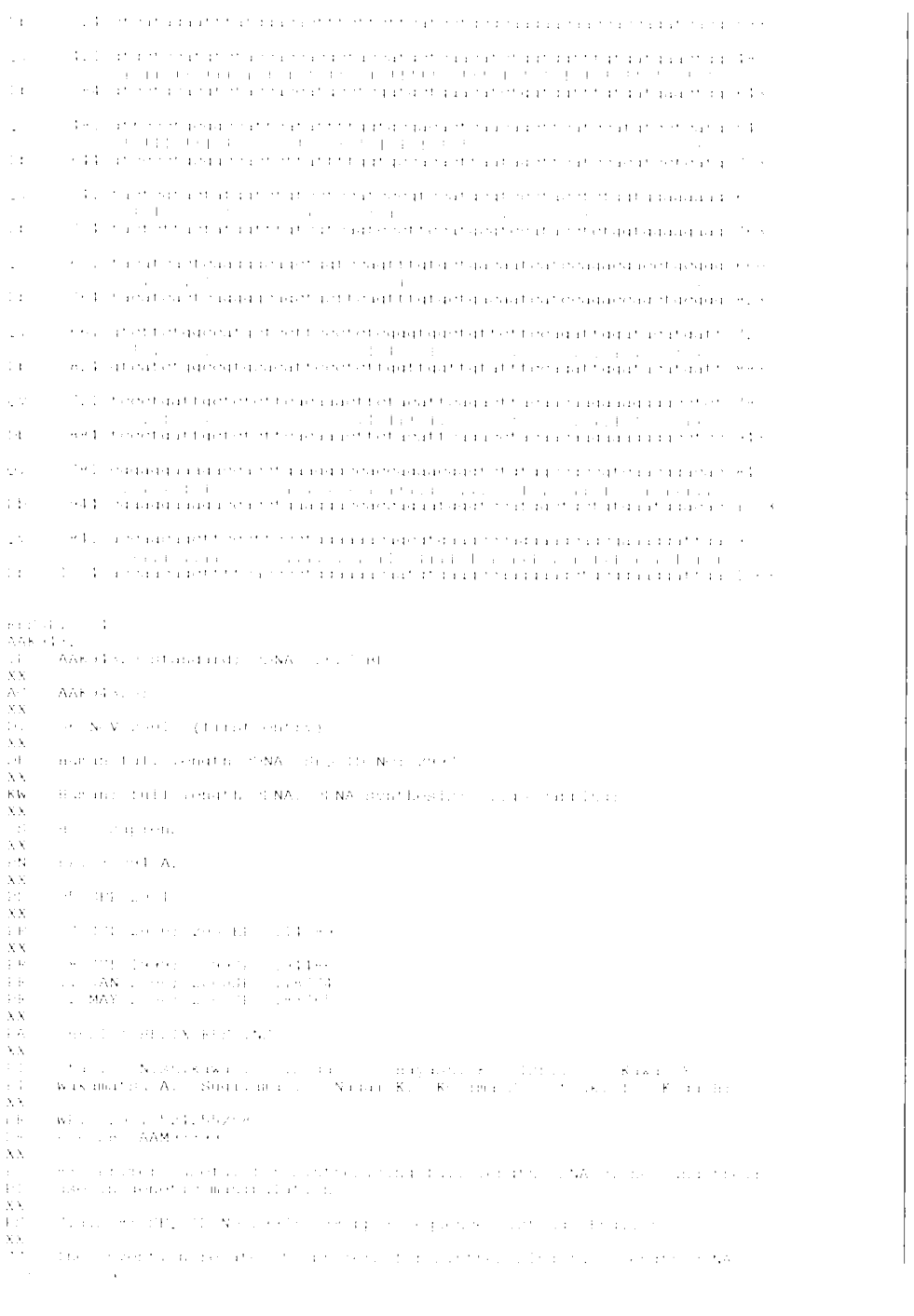


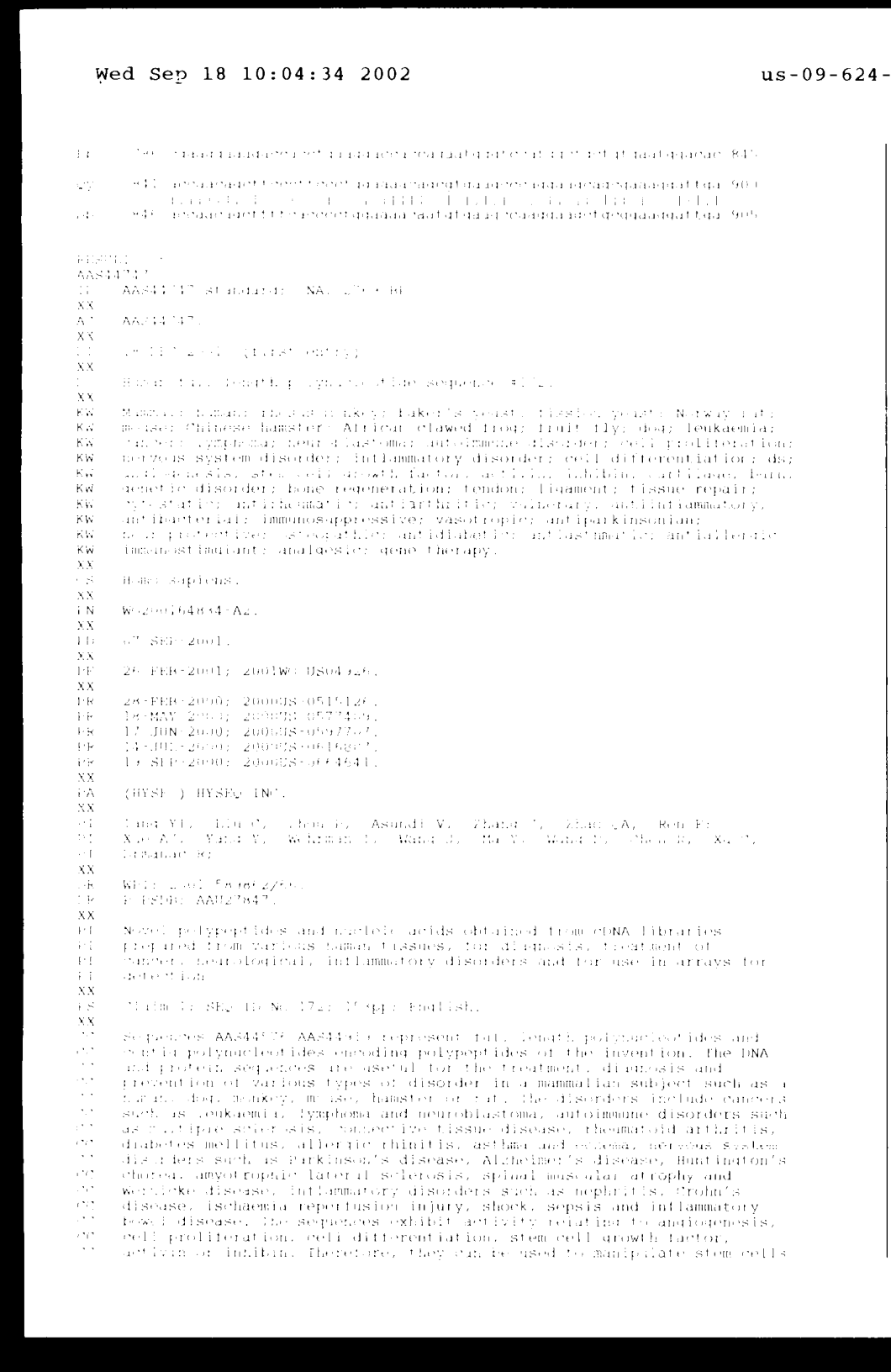


AAY64469
ID AAY64469 standard; cDNA; 2715 BP.
XX
AC AAY64469;
XX
DT 27-DEC-2001 (first entry)
XX
DE Human vesicle associated protein 1b encoding cDNA SEQ ID NO:88.
XX
KW human vesicle associated protein; VEAS; diagnosis; nootropic;
KW immunosuppressive; antibacterial; antihelminthic; protozacide;
KW fungicide; anti-inflammatory; cytostatic; gene therapy; akinesia; AIDS;
KW transport disorder; Bell's palsy; diabetes mellitus; catatonia;
KW infection; myositis; dementia; Grave's disease; autoimmune disorder;
KW inflammatory disorder; asthma; bronchitis; Reiter's syndrome; cancer;
KW systemic lupus erythematosus; rheumatoid arthritis; infection;
KW fungicide; 88.
XX
OS Homo sapiens.
XX



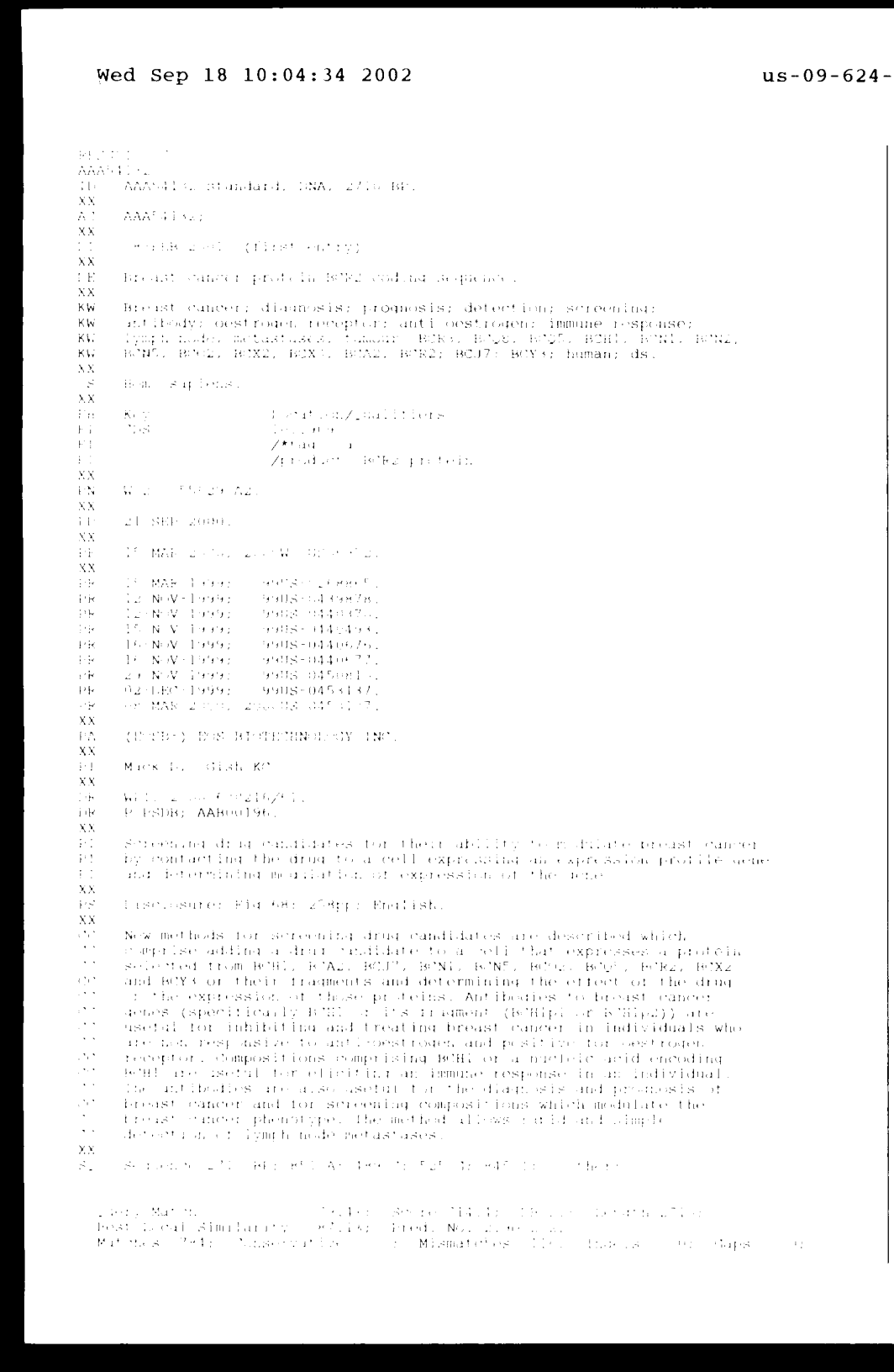
RN [2]
XX
RP 1-NOV-2000.
XX
RP 1-APR-2000; 2000WO-IB000454.
XX
PR 2-APR-1999; 99US-0127401.
PR 20-JUL-1999; 99US-0144701.
XX
PA (INCY) INCYTE PHARM INC.
XX
PI Lal P, Yue H, Hillman JL, Baughn MR, Tang YT, Gu HAM, Azimzai Y;
XX
DR WPI; 2000-665208/64.
DR P-PSDB; AAB24240.
XX
PT Novel vesicle associated proteins (VEAS) and polynucleotides encoding
PT VEAS for diagnosis, treating or preventing transport disorders (e.g.
PT Bell's palsy), autoimmune/inflammatory disorders (e.g. AIDS) or cancer
PT (e.g. leukemia)
XX
PS Claim 4; Page 111 112; 116pp; English.
XX
CC AAY60085 to AAY61000 and AAY64367 to AAY64469 encode the human vesicle
CC associated proteins 1 to 19 (VEAS-1 to 19) given in AAB24222 to AAB24240.
CC Human VEAS proteins have nootropic, immunosuppressive, antibacterial,
CC fungicide, antihelminthic, protozacide, fungicide, anti-inflammatory and
CC cytostatic activities. The isolated polypeptides and polynucleotides are
CC useful for diagnosis, treating or preventing transport disorders (e.g.
CC akinesia, Bell's palsy, diabetes mellitus, catatonia, infectious
CC myositis, dementia, Grave's disease or AIDS), autoimmune/inflammatory
CC disorders (e.g. AIDS, asthma, bronchitis, Reiter's syndrome, systemic
CC lupus erythematosus, rheumatoid arthritis, or viral, fungal, bacterial,
CC protozoal or helminthic infections), or cancer (e.g. adenocarcinoma,
CC leukemia, melanoma, sarcoma, or gall bladder, heart , kidney, bone,
CC brain, breast, pancreas or skin cancers). The polypeptides may also be
CC used to immunise a host animal. The polypeptides and polynucleotides are
CC further useful for screening libraries of compounds or for drug
CC screening assays.
XX
SQ Sequence 2715 BP; 691 A; 885 C; 486 G; 752 T; 0 other;

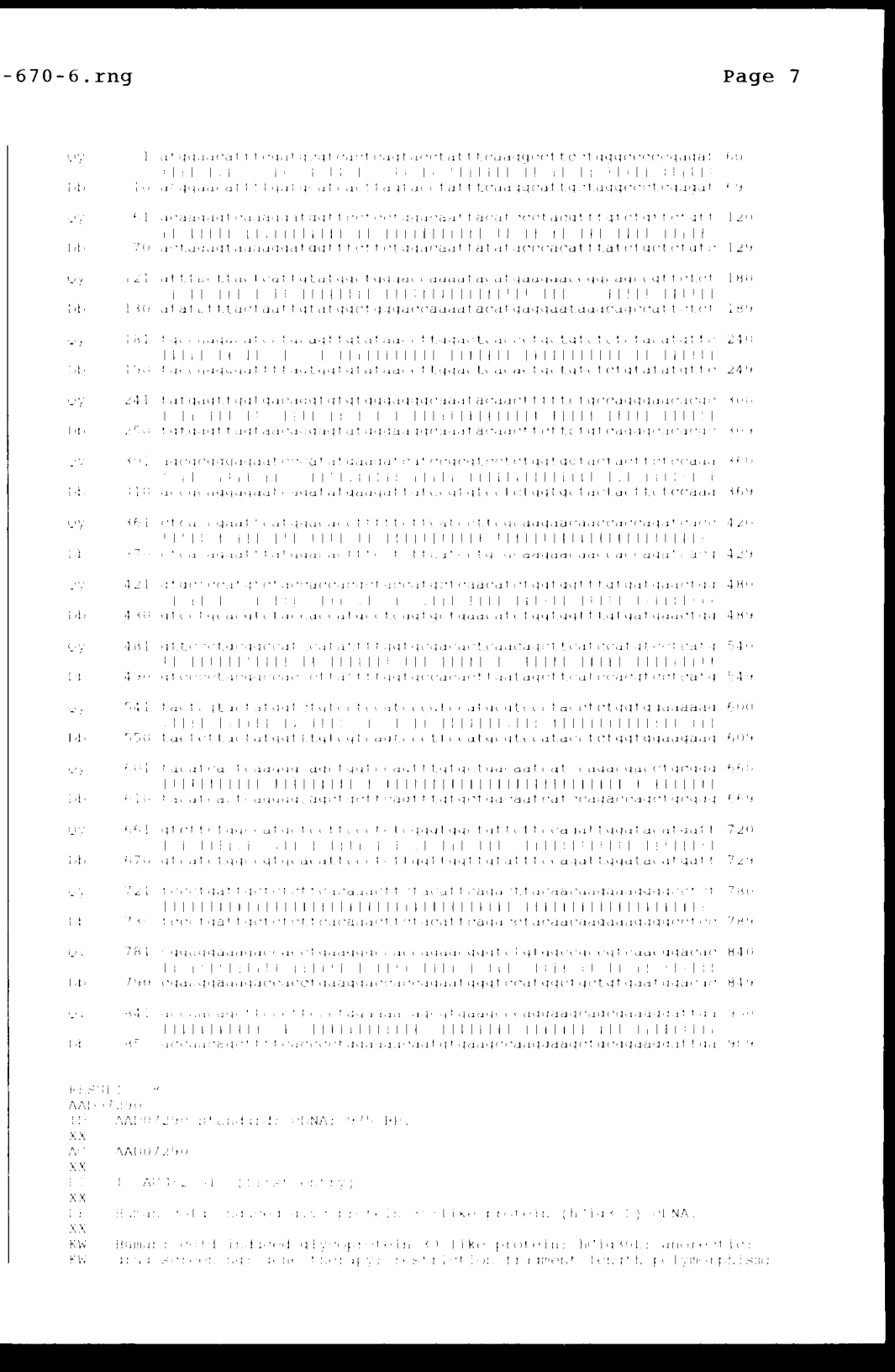




ID AAS44734 standard; DNA; 2769 BP.
XX
AC AAS44734;
XX
DT 26-FEB-2001 (first entry)
XX
DE Human bone marrow derived mesenchymal stem cell polynucleotide sequence #12.
XX
KW Mammalian; human; recombinant; baker's yeast; tissue; yeast; Norway rat;
KW mouse; Chinese hamster; African clawed frog; fruit fly; dog; leukaemia;
KW cancer; cytoma; sarcoma; autoimmune disorder; cell proliferation;
KW nervous system disorder; inflammatory disorder; cell differentiation; ds;
KW cell necrosis; stem cell growth factor; actin; inhibin; cartilage; burn;
KW genetic disorder; bone regeneration; tendon; ligament; tissue repair;
KW cytostatic; antirheumatic; antiarthritic; cytotoxic; antifibrinolytic;
KW antibacterial; immunosuppressive; vasotropic; antiparkinsonian;
KW wound protective; steroplasis; antidiabetic; antiasthmatic; antiallergic;
KW immunostimulant; analgesic; gene therapy.
XX
OS Homo sapiens.
XX
PN WO200164834-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US006126.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 16-MAY-2000; 2000US-0571459.
PR 17-JUN-2000; 2000US-0597727.
PR 13-JUL-2000; 2000US-0616982.
PR 15-SEP-2000; 2000US-0663641.
XX
PA (HYSE) HYSEQ INC.
XX
PI Tang Y.T., Liu C., Zhou P., Asundi V., Zhang Y., Zhao A., Ren F.,
PI Xie A.C., Yang Y., Wehrman T., Wang J., Ma Y., Wang E., Chen R., Xu C.,
PI Emanuel R.;
XX
DR WPI; 2001504826/54.
DR P-PSDB; AAU27847.
XX
PT Novel polypeptides and nucleic acids obtained from cDNA libraries
PT prepared from various human tissues, for diagnosis, treatment of
PT cancer, neurological, inflammatory disorders and for use in arrays for
PT detection.
XX
PS Claim 1; SEQ ID NO 1722; 1729pp; English.
XX
CC Sequences AAS44673-AAS44815 represent full length polynucleotides and
CC partial polynucleotides encoding polypeptides of the invention. The DNA
CC and protein sequences are useful for the treatment, diagnosis and
CC prevention of various types of disorder in a mammalian subject such as a
CC human, dog, monkey, mouse, hamster or rat. The disorders include cancers
CC such as leukaemia, lymphoma and neuroblastoma, autoimmune disorders such
CC as multiple sclerosis, connective tissue disease, rheumatoid arthritis,
CC diabetes mellitus, allergic rhinitis, asthma and eczema, nervous system
CC disorders such as Parkinson's disease, Alzheimer's disease, Huntington's
CC chorea, amyotrophic lateral sclerosis, spinal muscular atrophy and
CC Wernicke disease, inflammatory disorders such as nephritis, Crohn's
CC disease, ischaemia reperfusion injury, shock, sepsis and inflammatory
CC bowel disease. The sequences exhibit activity relating to angiogenesis,
CC cell proliferation, cell differentiation, stem cell growth factor,
CC action or inhibin. Therefore, they can be used to manipulate stem cells
CC in culture to give rise to neuroepithelial cells that can be used to
CC augment or replace cells damaged by illness, accidental damage or genetic
CC disorders. The sequences may also be used for regeneration of bone,
CC cartilage, tendons and ligaments and in tissue repair and burn healing.
CC Note: Some sequences for this patent did not form part of the printed
CC specification, but were obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2769 BP; 799 A; 543 C; 554 G; 873 T; 0 other;

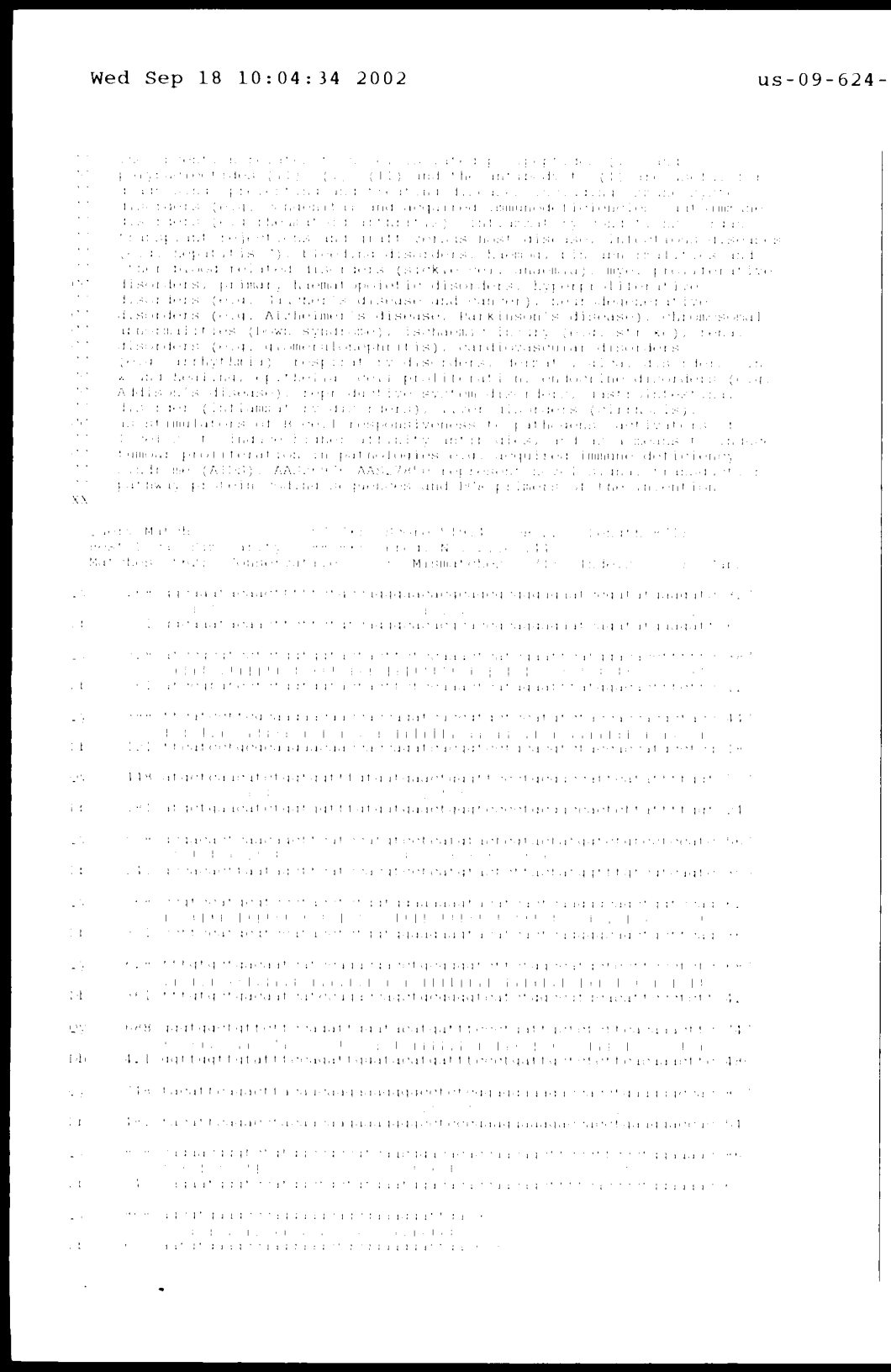
Query Match 17.4%; Score 726; DB 22; Length 2769;
Best Local Similarity 87.2%; Pred. No. 9.8e-203;
Matches 745; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

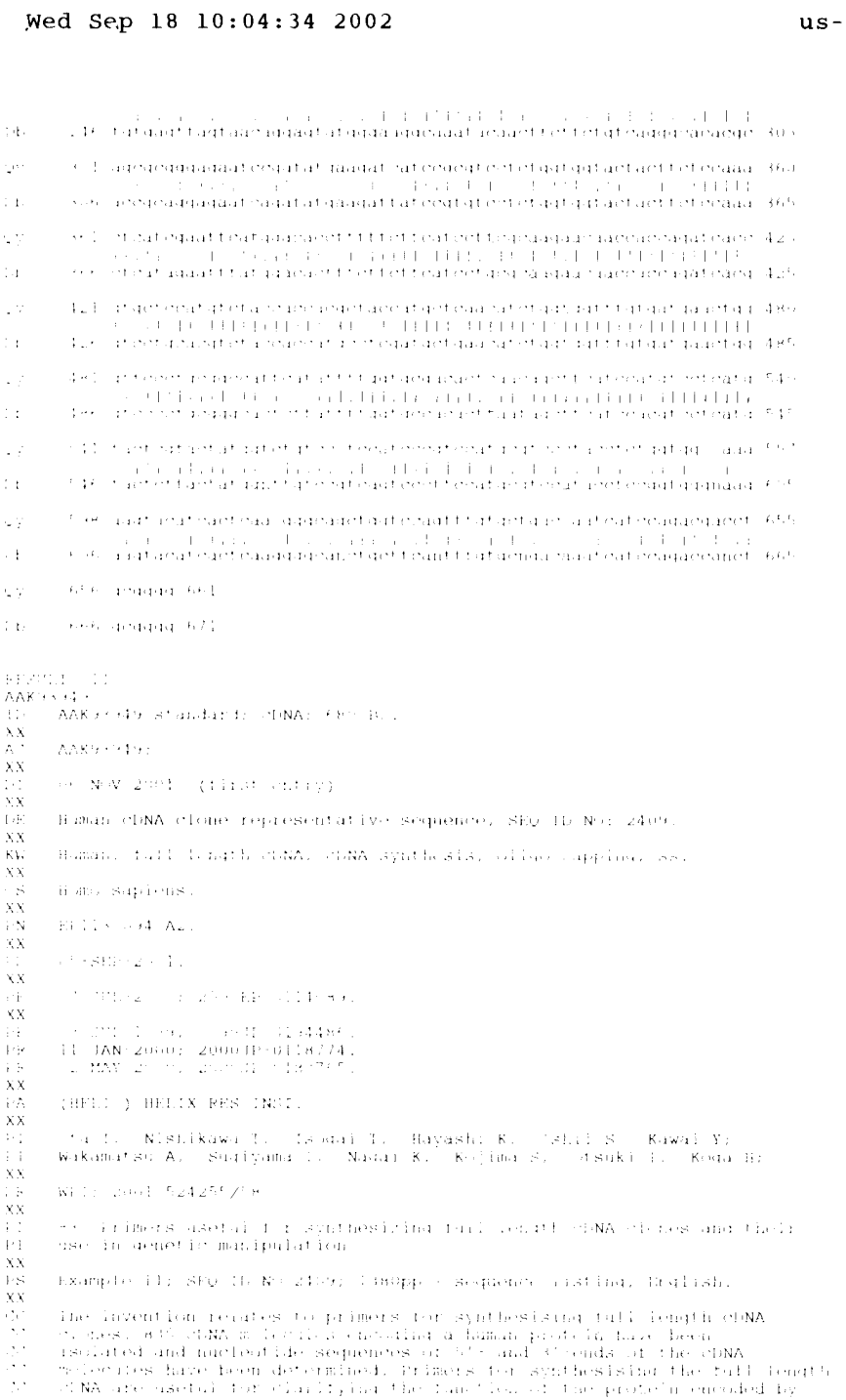




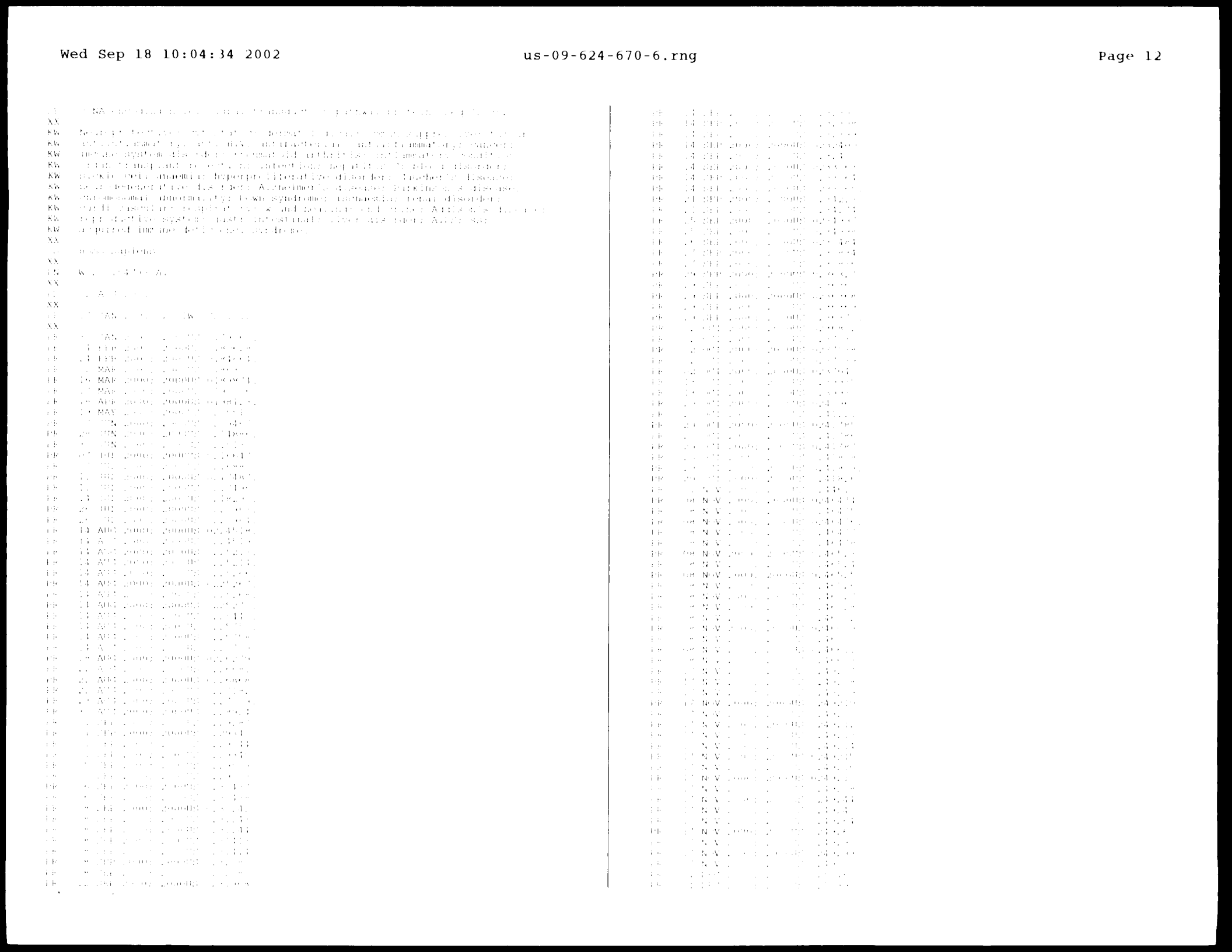
PR 16-MAR-2000; 2000US-0116078.
PR 17-MAR-2000; 2000US-0116078.
PR 17-APR-2000; 2000US-0119331.
PR 17-MAY-2000; 2000US-0120515.
PR 30-JUN-2000; 2000US-0121467.
PR 29-SEP-2000; 2000US-0214886.
PR 29-SEP-2000; 2000US-0215765.
PR 29-SEP-2000; 2000US-0216643.
PR 29-SEP-2000; 2000US-0216844.
PR 29-SEP-2000; 2000US-0217367.
PR 29-SEP-2000; 2000US-0217396.
PR 29-SEP-2000; 2000US-0218245.
PR 29-SEP-2000; 2000US-0220063.
PR 29-SEP-2000; 2000US-0220064.
PR 01-AUG-2000; 2000US-0224518.
PR 03-AUG-2000; 2000US-0224714.
PR 03-AUG-2000; 2000US-0225213.
PR 03-AUG-2000; 2000US-0225214.
PR 03-AUG-2000; 2000US-0225906.
PR 03-AUG-2000; 2000US-0225907.
PR 14-AUG-2000; 2000US-0225908.
PR 03-AUG-2000; 2000US-0225917.
PR 08-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225707.
PR 14-AUG-2000; 2000US-0225708.
PR 03-AUG-2000; 2000US-0225709.
PR 14-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226741.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227002.
PR 29-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229445.
PR 05-SEP-2000; 2000US-0229514.
PR 05-SEP-2000; 2000US-0230637.
PR 06-SEP-2000; 2000US-0230638.
PR 07-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231313.
PR 08-SEP-2000; 2000US-0231314.
PR 08-SEP-2000; 2000US-0231315.
PR 12-SEP-2000; 2000US-0232008.
PR 13-SEP-2000; 2000US-0232407.
PR 13-SEP-2000; 2000US-0232408.
PR 13-SEP-2000; 2000US-0232409.
PR 13-SEP-2000; 2000US-0232501.
PR 13-SEP-2000; 2000US-0232502.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234774.
PR 25-SEP-2000; 2000US-0234997.
PR 27-SEP-2000; 2000US-0235682.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235804.
PR 27-SEP-2000; 2000US-0235805.
PR 27-SEP-2000; 2000US-0235807.
PR 28-SEP-2000; 2000US-0236868.
PR 28-SEP-2000; 2000US-0236964.
PR 28-SEP-2000; 2000US-0236967.
PR 29-SEP-2000; 2000US-0237042.
PR 29-SEP-2000; 2000US-0237167.
PR 29-SEP-2000; 2000US-0237168.
PR 29-SEP-2000; 2000US-0237171.

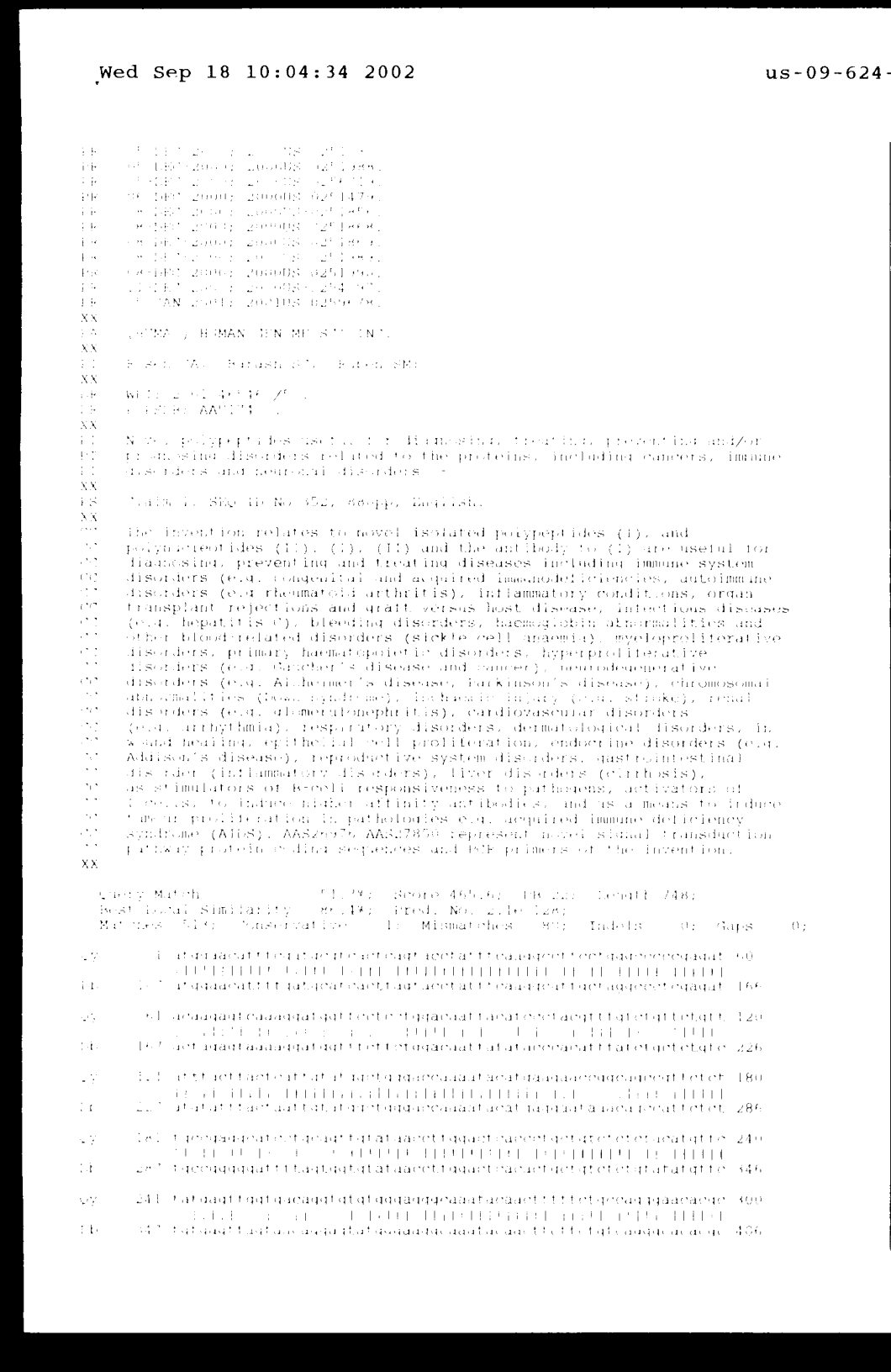
PR 10-OCT-2000; 2000US-0240044.
PR 13-OCT-2000; 2000US-0240947.
PR 27-OCT-2000; 2000US-0241095.
PR 30-OCT-2000; 2000US-0241221.
PR 27-OCT-2000; 2000US-0241785.
PR 27-OCT-2000; 2000US-0241786.
PR 27-OCT-2000; 2000US-0241787.
PR 27-OCT-2000; 2000US-0241808.
PR 27-OCT-2000; 2000US-0241809.
PR 27-OCT-2000; 2000US-0241825.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 09-NOV-2000; 2000US-0246611.
PR 03-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249243.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249263.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249269.
PR 17-NOV-2000; 2000US-0249270.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250161.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251698.
PR 05-DEC-2000; 2000US-0251719.
PR 06-DEC-2000; 2000US-0251879.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251878.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251880.
PR 08-DEC-2000; 2000US-0251900.
PR 11-DEC-2000; 2000US-0254067.
PR 05-JAN-2001; 2001US-0256726.
XX
PA (HUMA) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Baker KP, Haldeman BA;
XX
PI WO; 200064923-A1.
PI US; AAU74401.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT treating diseases related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
PS Claim 1; SEQ ID No 654; 880pp; English.
XX

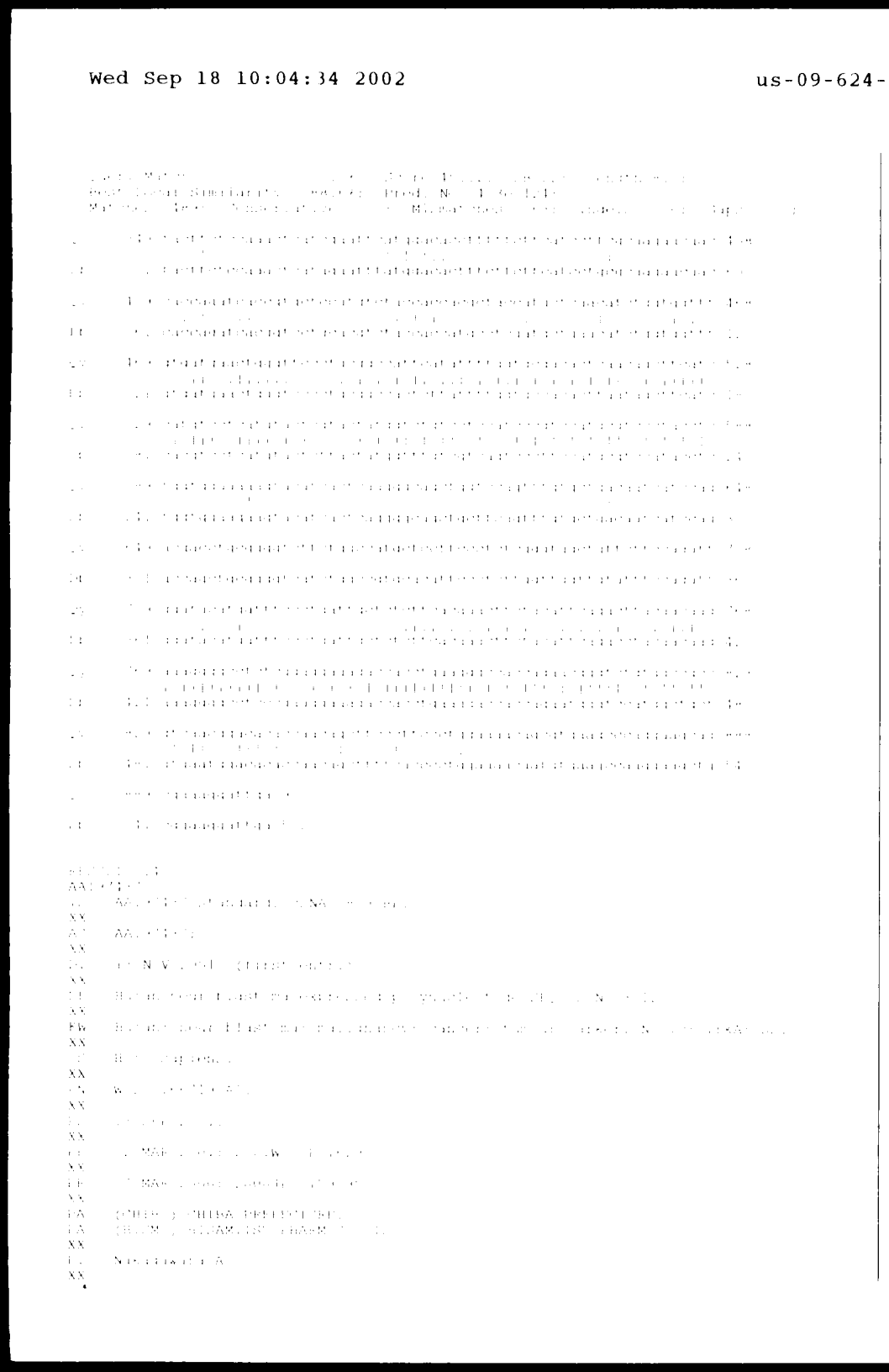


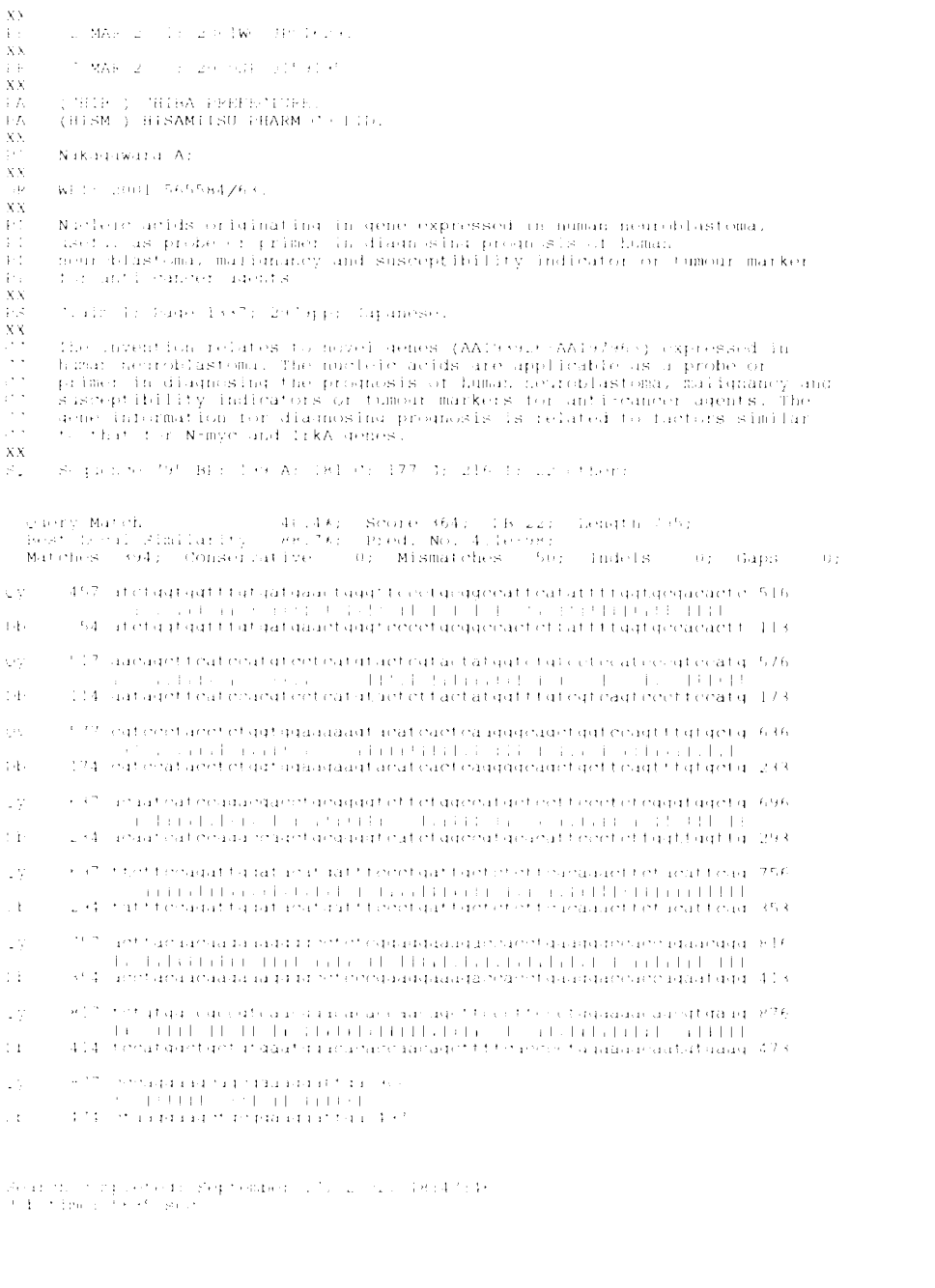


```
RESULT 12
AAK24049
ID   AAK24049 standard; cDNA; 690 BP.
XX
AC   AAK24049;
XX
DT   29-NOV-2001  (first entry)
XX
DE   Human cDNA clone representative sequence, SEQ ID NO: 2409.
XX
KW   Human; full length cDNA; cDNA synthesis; office application; ss.
XX
OS   Homo sapiens.
XX
PN   EP1154024-A2.
XX
PD   14-NOV-2001.
XX
PF   09-MAY-2001; 2001EP-0111274.
XX
PR   11-JAN-2000; 2000JP-0001274.
PR   09-MAY-2000; 2000JP-0136774.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T; Nishikawa T; Isogai T; Hayashi K; Ishii S; Kawai Y;
PI   Wakamatsu A; Sugiyama T; Nagai K; Kojima S; Otsuki T; Kona H;
XX
DR   WPI; 2001-524259/58.
XX
PT   New primers useful for synthesizing full length cDNA clones and their
PT   use in genetic manipulation
XX
PS   Example 11; SEQ ID NO 2409; 134pp; Sequence listing, English.
XX
CC   The invention relates to primers for synthesizing full length cDNA
CC   clones. A cDNA library synthesized from a human protein have been
CC   isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC   fragments have been determined. Primers for synthesizing the full length
CC   cDNA are useful for clarifying the functions of the protein encoded by
CC   the cDNA. The full length cDNA clones were obtained by construction of full
CC   length enriched cDNA libraries that were synthesized by the oligo capping
CC   method. The primers enable the production of the full length cDNA easily
CC   without any special methods. The present sequence was used as the
CC   representative sequence from a human clone which was used in
CC   homology searches to identify the clones.
CC   Note: the sequence data for this patent did not form part of the printed
CC   specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ   Sequence 690 BP; 167 A; 174 C; 144 G; 199 T; 6 other;


  Query Match          92.8%;  Score 475;  DB 22;  Length 690;
  Best Local Similarity   95.7%;  Pred. No. 3.2e-132;
  Matches  571;  Conservative    0;  Mismatches   26;  Indels    5;  Gaps    4;

Qy       1 atgaaaaattttaaaatcaatctatttattcaaaccttctaaaccccaaatat 50
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 atgaaaaattttaaaatcaatctatttattcaaaccttctaaaccccaaatat 50

Qy      51 ttaaaaatcaaaatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      51 ttaaaaatcaaaatcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 100

Qy     101 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 150
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     101 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 150

Qy     151 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     151 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 200

Qy     201 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 250
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     201 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 250

Qy     251 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     251 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 300

Qy     301 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 350
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 350

Qy     351 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     351 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 400

Qy     401 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 450
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     401 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 450

Qy     451 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     451 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 500

Qy     501 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 550
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     501 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 550

Qy     551 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     551 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 600

Qy     601 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 650
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 650

Qy     651 aaaaaaaaaa 660
          ||||||||||
Db     651 aaaaaaaaaa 660


RESULT 13
AAS27417
ID   AAS27417 standard; cDNA; 748 BP.
XX
AC   AAS27417;
XX
DT   07-NOV-2001  (first entry)
XX
```









GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2002, 17:13:51 ; Search time 290.2 Seconds
(without alignments)
5260.440 Million cell updates/sec

Title: US-09-624-670-5
Perfect score: 87
Sequence: 1 [illegible] .... [illegible] 87

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736648 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: [illegible]

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_[illegible]2802:*
1: [illegible]*
2: [illegible]*
3: [illegible]*
4: [illegible]*
5: [illegible]*
6: [illegible]*
7: [illegible]*
8: [illegible]*
9: [illegible]*
10: [illegible]*
11: [illegible]*
12: [illegible]NA1991.DAT:*
13: [illegible]*
14: [illegible]*
15: [illegible]*
16: [illegible]*
17: [illegible]*
18: [illegible]*
19: [illegible]*
20: [illegible]*
21: [illegible]*
22: [illegible]*
23: [illegible]*
24: [illegible]*

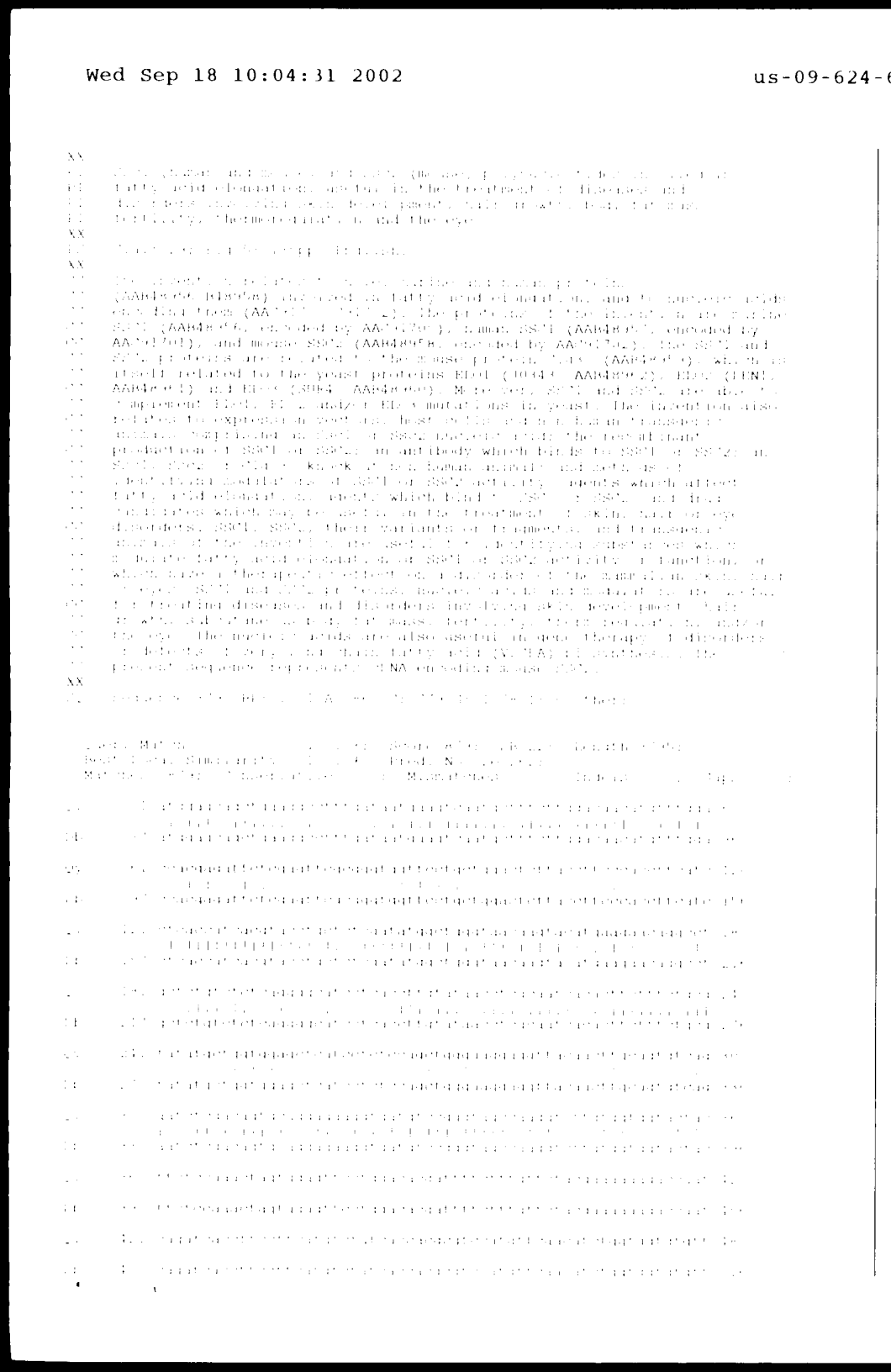
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

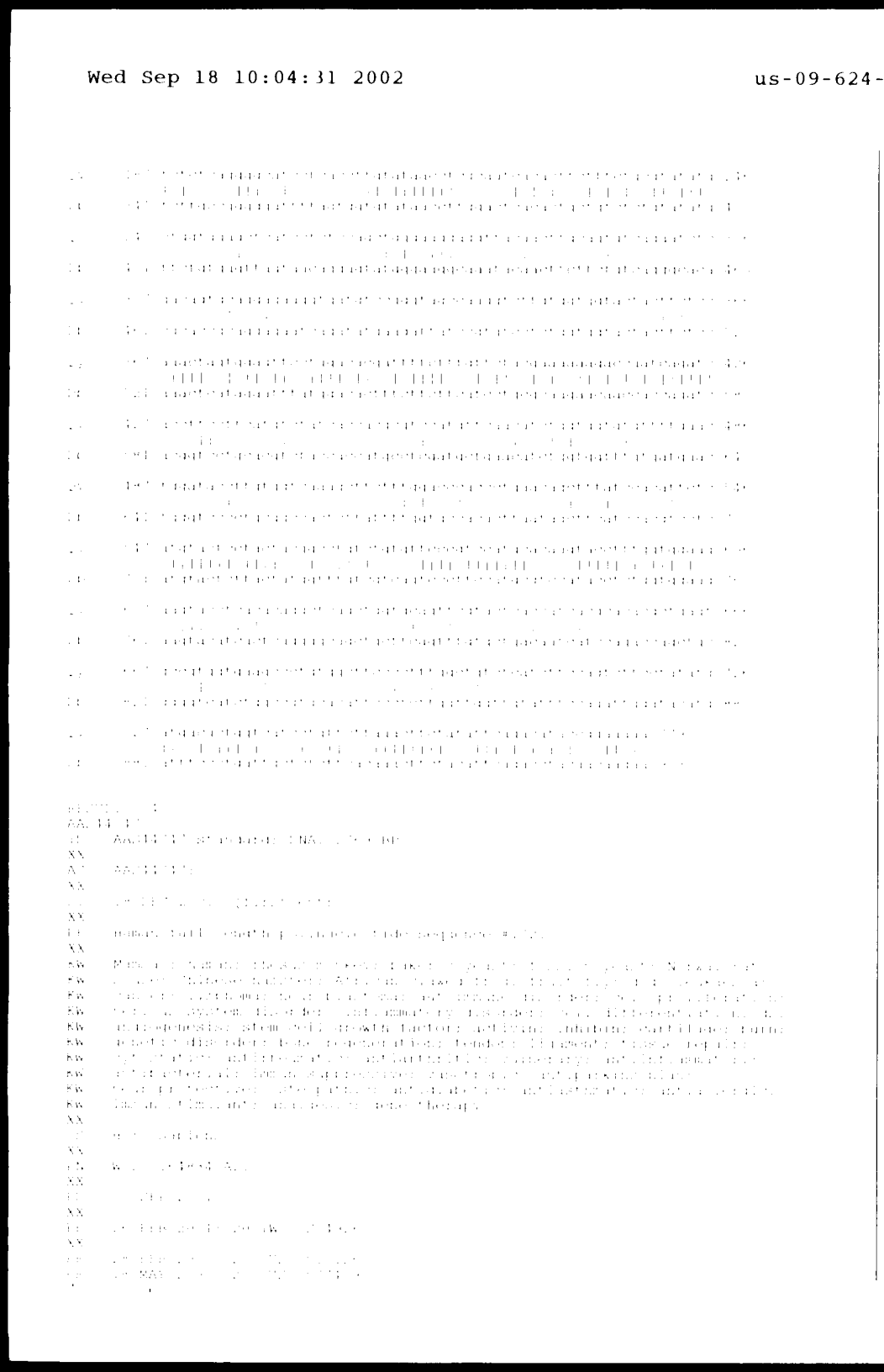
Result No.	Score	% Query Match	Length	DB	ID	Description
1	87	100.0	[illegible]	22	AAC[illegible]	Mouse fatty acid e
2	333	37.9	2627	22	AAK[illegible]	Human full-length
3	331.4	37.7	2313	21	AAC64569	Human vesicle asso
4	331.4	37.7	2764	22	AAS14747	Human full-length
5	331.4	37.7	3083	22	AAS14919	Human contig polyn
6	330.6	37.6	914	21	AAZ94211	Human elongase HSE
7	329.8	37.5	2710	21	AAA[illegible]132	Breast cancer prot
8	283	32.2	25[illegible]	[illegible]	AAN81158	Fragment from cl[illegible]
9	258	29.4	748	22	AAS27317	cDNA encoding nove
10	253	28.8	680	22	AAK91344	Human cDNA 5' end
11	253	28.8	680	22	AAK93549	Human cDNA clone r
12	246.6	28.1	975	22	AA[illegible]	Human cold-induced
13	241	27.4	871	22	AAS27318	cDNA encoding nove
14	214	24.3	817	22	AAK93682	Human cDNA clone r
15	177.6	20.2	1283	23	AAS91555	DNA encoding novel
16	156	17.7	809	22	AAI97437	Human neuroblastom
17	152.4	17.3	795	22	AAI95727	Human neuroblastom
18	148.4	16.9	927	22	AAK93155	Human cDNA clone r
19	136.8	15.6	838	23	ABL09165	Drosophila melanog
20	135.2	15.4	831	23	ABL05185	Drosophila melanog
21	134.8	15.3	3734	23	ABL05184	Drosophila melanog
22	132.4	15.1	24664	22	AAK91310	Human digestive sy
23	132.4	15.1	24664	22	AAS32137	Human liver associ
24	131.4	14.9	2225	21	AAC77654	Human cancer assoc
25	127.8	14.5	1914	23	ABL15699	Drosophila melanog
26	127	14.4	1233	23	ABL11961	Drosophila melanog
27	127	14.4	4336	23	ABL11960	Drosophila melanog
28	117.6	13.4	488	23	AAS76145	DNA encoding novel
29	117.2	13.3	1329	23	ABL20787	Drosophila melanog
30	112.6	12.8	3302	23	ABL[illegible]164	Drosophila melanog
31	111.6	12.7	549	21	AAC95232	Cat flea head and
32	110.2	12.5	1583	23	ABL[illegible]223	Drosophila melanog
33	108.4	12.3	874	22	AAK91805	Human cDNA 5' end
34	108.4	12.3	874	22	AAK93947	Human cDNA clone r
35	108.4	12.3	2190	22	AAK94230	Human full-length
36	105.8	12.2	1457	22	AAH14551	Human cDNA sequence
37	105.8	12.0	1482	19	AAV59752	Human secreted pro
38	105.8	12.0	1542	19	AAV59636	Human secreted pro
39	105.8	12.0	1566	21	AAZ56736	Human transmembran
40	[illegible]	[illegible]	[illegible]	22	AAC[illegible]	Human fatty acid e
41	[illegible]	[illegible]	[illegible]	[illegible]	ABL[illegible]	Drosophila melanog
42	[illegible]	[illegible]	[illegible]	22	AAC[illegible]	Mouse fatty acid e
43	[illegible]	[illegible]	[illegible]	22	AAA08806	Human secreted pro
44	[illegible]	[illegible]	[illegible]	[illegible]	AAA[illegible]123	Mortierella alpina
45	[illegible]	[illegible]	[illegible]	22	AAL48732	Human breast cell

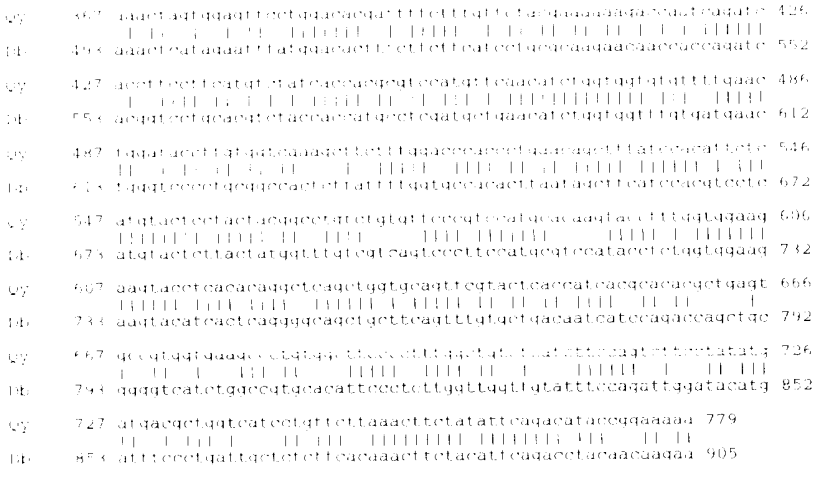
ALIGNMENTS

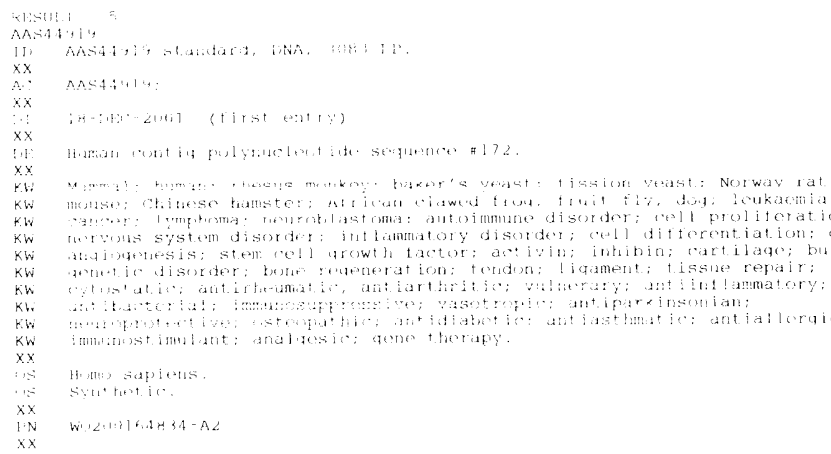
RESULT 1
AAC[illegible]
ID AAC[illegible] standard; DNA; [illegible] BP.
XX
AC AAC[illegible];
XX
DT [illegible] (first entry)
XX
DE [illegible] fatty acid [illegible] DNA.
XX
KW Mouse SSC2; murine; fatty acid elongation; very long chain fatty acid;
KW [illegible] knockout animal;
KW skin development; hair growth; subcutaneous body fat; fertility;
KW thermoregulation; dermatological; ophthalmological; antiinfertility;
KW drug screening; gene therapy; ss.
XX
OS Mus sp.
XX
PN WO200070945-A2.
XX
PD 30-NOV-2000.
XX
PF 16-MAY-2000; 2000WO-EP04371.
XX
PR 20-MAY-1999; 99US-0135084.
XX
PA (KAR-) KAROLINSKA INNOVATIONS AB.
XX
PI Jacobsson A, Asadi A, Westerberg R, Nedergaard J, Tvrdik P;
XX
DR WPI; 2001-016322/02.
DR P-PSDB; AAB48[illegible].

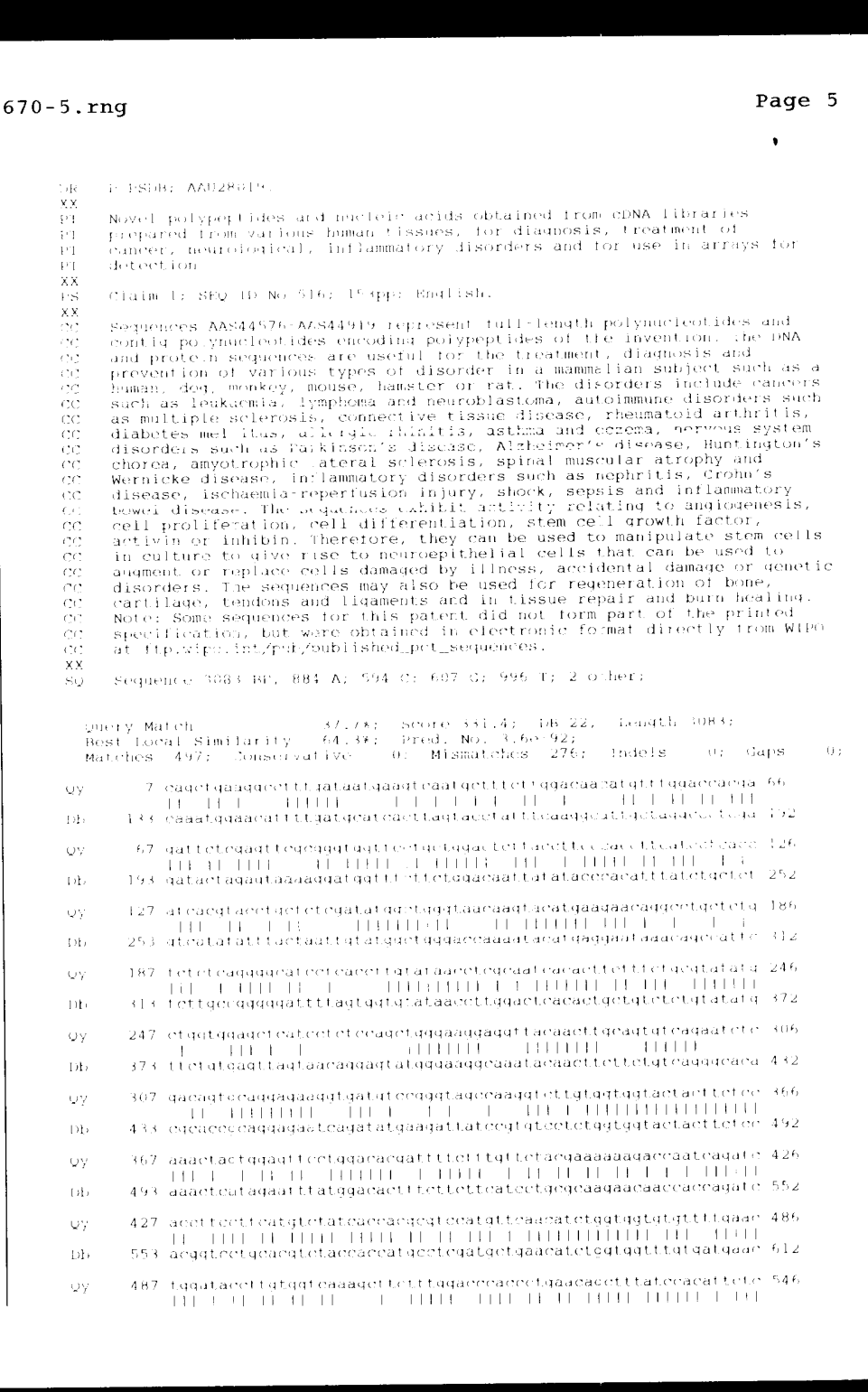


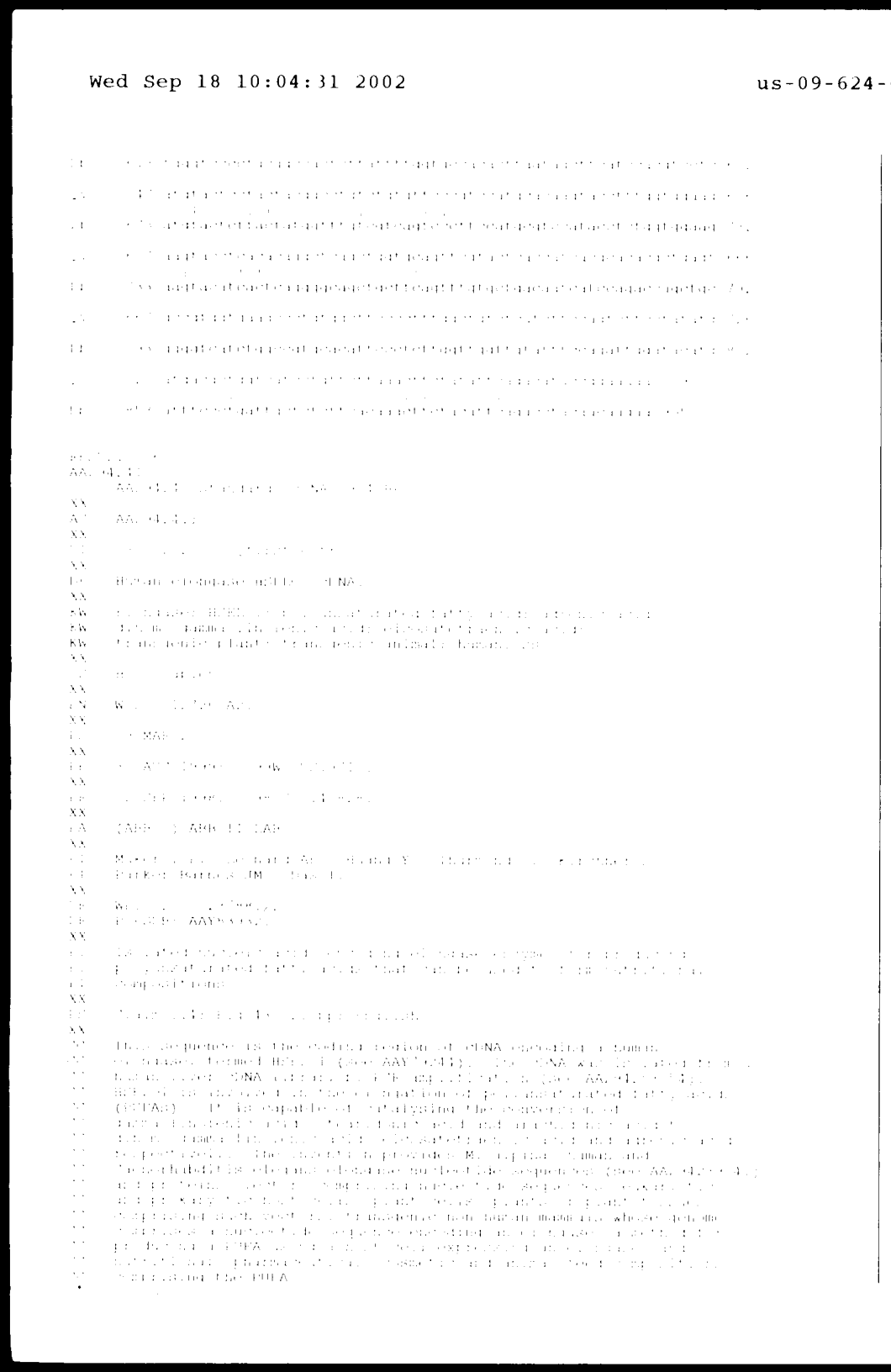
```
XX   length enriched cDNA libraries that were synthesised by the oligo-capping
XX   method. The primers enable the production of the full length cDNA easily
XX   without any special methods. The present sequence is a full length
XX   human cDNA of the invention.
XX   Note: the sequence data for this patent did not form part of the printed
XX   specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ   Sequence 2627 BP; 771 A; 490 C; 520 G; 846 T; 0 other;


  Query Match             57.9%;  Score 1113;  DB 21;  Length 2627;
  Best Local Similarity   64.4%;  Pred. No. 1e-92;
  Matches  496;  Conservative    0;  Mismatches  275;  Indels    0;  Gaps    0;

Qy        7 caqctqaaaqcctttqataataaqtcaatqctttcttqaacaacatatttqqacccacca  66
            ||   | |    ||||||     | | | | | | ||  |     || |  | || |||
Db        5 caaatqqaaaratttqatqcatcacttaqtacctatttcaaaqcattqctaqqcctcqa  62

Qy       67 qattctcqaqttcqcqqatqattcctqctqqactcttaccttcccaccttcatcctcacc  126
            ||| || ||||    || |||  | ||||   |||  | ||| | ||  ||| |  |
Db       63 qatactaqaatcaaaqqatqattttcttctqqacaattatatacccacatttatctqctct  122

Qy      127 atcacqtaccqtqctctqatatqqctqqctaacaaqtacatqaaaaacaqqcctcctctq  186
             ||| ||  | ||     |||| ||||    |  ||||||| ||| |  :   || |
Db      123 qtcatatatttactaattatatqqctqqqacctaaatacatqaqqaataaacaqcctttc  182

Qy      187 tctctcaqqqqcatcctcaccttqtataacctcqcaatcacacttctttctacqtatatq  246
            ||| : |||| || :     ||||||||| | | ||||| || |||  |||||||
Db      183 tctqcc:qqqqatttttatqatqtataacctcqqactcacactqtatctctqtatata  242

Qy      247 ctqqtqaaactcatcctctccaqctqqaaqaqaqqttacaacttqcaqtqtcaqaaatctc  306
             :   ||| |  |              ||||||||    ||||||||    ||:|||
Db      243 ttctqtaqttaataaraqaaatarqqaaaqaaaaatacaacttctctqtcaqqqcaca  302

Qy      307 aacaqtqcaqqaaaaaqqtaatqtccqqataqccaaqqtcttqtqqtqqtactactcttcc  366
             ||  |||||||||    |'| |  | |    |  ||| | ||||||||||||||||||
Db      303 cacaccqaaqaqqaaatcaqatataqaaqattatccqtqtcctctqatqatactacttctcc  362

Qy      367 aaactatqqaattctqtaacaqatttctttctatctacqaaaaaqaaccaatcaqatc  426
            |||||  | | | |  ||||||  | |||||  | || || || || | | | ||||||
Db      363 aaactcatatqaatttatqcacactttcttcttcatctqcacaqaqacaacaccaqatc  422

Qy      427 acctcctttatatctatcaccaqcaqtccatattcaaacatcqatqatqtatttqaaac  486
            ||  |||| || |||| |||| | || || ||| | ||||||||||| | |  |||||
Db      423 acqqtcctqcacqtctacqacatqcctqatctqaacatcaatqatttatcaaqaac  482

Qy      487 tqqataccttqtqqtcaaaacttctttqqaccaccqtqaacaqqcttatcccaqttctc  546
            ||: | || || || ||    | |||||  |||| || || |||| ||| || | |||
Db      483 tqqqtcccctqcqqccactcttatttqqtqccacacttaatqqcttcatccacqtcctc  542

Qy      547 atqtactcctactacqqcctqtctqtattcccqtccatqcacaaqtaccttqqtqqaaq  606
            |||||||| |||| ||  |||    |||| ||||||     |||| ||||||||
Db      543 atqtactcttactatcqqtttatcqtcaqtcccttccatacqtccatacctctqqtqqaaq  602

Qy      607 aaqtacctcacacaqactcqqctqqtqcaqttcqtactcaccatcacqcacacqctqaqt  666
            ||||| ||| :|:|  |||||| | ||||| || || ||||  || ||     |
Db      603 aaqtacatcacactcaqaqqcaqctqcttcaattatqctqacaatatccaqaccaqctqc  662

Qy      667 qtcqtqtqtqaaqcctqtaqcttcccttqqctqtctcatttcaatttcctatata  726
             | ||  | ||| ||    ||||  |||| || |  ||||| |  |||| |||||
Db      663 qaqatcatctaccctqcaqcattcccttcttqqttaqttatttcaaattqaatacatq  722

Qy      727 atqcacqtqqtcatcatctqttcttaaacttctatatttcaaacatacccqaaaaa  779
            |  |  |:|:|    || |||  |||||||| ||||||  |||  ||  ||
Db      723 atttcctcaattqctctcttcaaaacttctacatttcaaacatacaacaaaaa  775


RESULT   5
AAC64369
ID   AAC64369 standard; cDNA; 2313 BP.
XX
AC   AAC64369;
XX
DT   17-FEB-2001  (first entry)
XX
DE   Human vesicle associated protein 19 encoding cDNA SEQ ID NO:38.
XX
KW   Human; vesicle associated protein; VEAS; diagnosis; nootropic;
KW   immunosuppressive; antibacterial; antihelminthic; protozoacide;
KW   virucide; anti inflammatory; cytostatic; gene therapy; akinesia; AIDS;
KW   transport disorder; Bell's palsy; diabetes mellitus; catatonia;
KW   infectious myositis; dementia; Grave's disease; autoimmune disorder;
KW   inflammatory disorder; asthma; bronchitis; Reiter's syndrome; cancer;
KW   systemic lupus erythematosus; rheumatoid arthritis; infection;
KW   tunqicide; ss.
XX
OS   Homo sapiens.
XX
PN   WO200060082-A2.
XX
PD   12-OCT-2000.
XX
PF   06-APR-2000; 2000WO-US09453.
XX
PR   07-APR-1999;    99US-0128193.
PR   20-JUL-1999;    99US-0144761.
XX
PA   (INCY ) INCYTE GENOMICS INC.
XX
PI   Lal P, Yue H, Hillman JL, Baughn MR, Tang YT, Lu DAM, Azimzai Y;
XX
DR   WPI; 2000-665008/64.
DR   P-PSDB; AAB24240.
XX
PT   New vesicle associated proteins (VEAS) and polynucleotides encoding
PT   VEAS for diagnosing, treating or preventing transport disorders (e.g.
PT   Bell's palsy), autoimmune/inflammatory disorders (e.g. AIDS) or cancer
PT   (e.g. leukemia)
XX
PS   Claim 4; Page 111-112; 112pp; English.
XX
CC   AAC59985 to AAC61600 and AAC64367 to AAC64369 encode the human vesicle
CC   associated proteins 1 to 19 (VEAS-1 to 19) given in AAB24222 to AAB24240.
CC   Human VEAS proteins have neotropic, immunosuppressive, antibacterial,
CC   tunqicide, antihelminthic, protozoacide, virucide, anti-inflammatory and
CC   cytostatic activities. The isolated polypeptides and polynucleotides are
CC   useful for diagnosing, treating or preventing transport disorders (e.g.
CC   akinesia, Bell's palsy, diabetes mellitus, catatonia, infectious
CC   myositis, dementia, Grave's disease or AIDS), autoimmune/inflammatory
CC   disorders (e.g. AIDS, asthma, bronchitis, Reiter's syndrome, systemic
CC   lupus erythematosus, rheumatoid arthritis, or viral, fungal, bacterial,
CC   protozoal or helminthic infections), or cancer (e.g. adenocarcinoma,
CC   leukaemia, melanoma, sarcoma, or gall bladder, heart , kidney, bone,
CC   brain, breast, pancreas or skin cancers). The polypeptides may also be
CC   used to immunise a host animal. The polypeptides and polynucleotides are
CC   further useful for screening libraries of compounds or for drug
CC   screening assays.
XX
SQ   Sequence 2313 BP; 645 A; 486 C; 480 G; 702 T; 0 other;


  Query Match             57.7%;  Score 1111;  DB 21;  Length 2313;
  Best Local Similarity   64.4%;  Pred. No. 3.1e-92;
  Matches  497;  Conservative    0;  Mismatches  276;  Indels    0;  Gaps    0;

Qy        7 caqctqaaaqcctttqataataaqtcaatqctttcttqaacaacatatttqqacccacca  66
            ||   | |    ||||||     | | | | | | ||  |     || |  | || |||
Db      161 caaatqqaaacatttqatqcatcacttaqtacctatttcaaaqcattqctaqqcctcqa  220

Qy       67 qattctcqaqttcqcqqatqattcctqctqqactcttaccttcccaccttcatcctcacc  126
            ||| || ||||    || |||||  | ||||   |||  | ||| | ||  ||| |  |
Db      221 qatactaqaatcaaaqqatqattttcttctqqacaattatatacccacatttatctqctct  280

Qy      127 atcacqtaccqtqctctqatatqqctqqctaacaaqtacatqaaaaacaqqcctcctctq  186
             ||| ||  | ||     |||| ||||    |  ||||||| ||| |  :   || |
Db      281 qtcatatatttactaattatatqqctqqqacctaaatacatqaqqaataaacaqcctttc  340
```

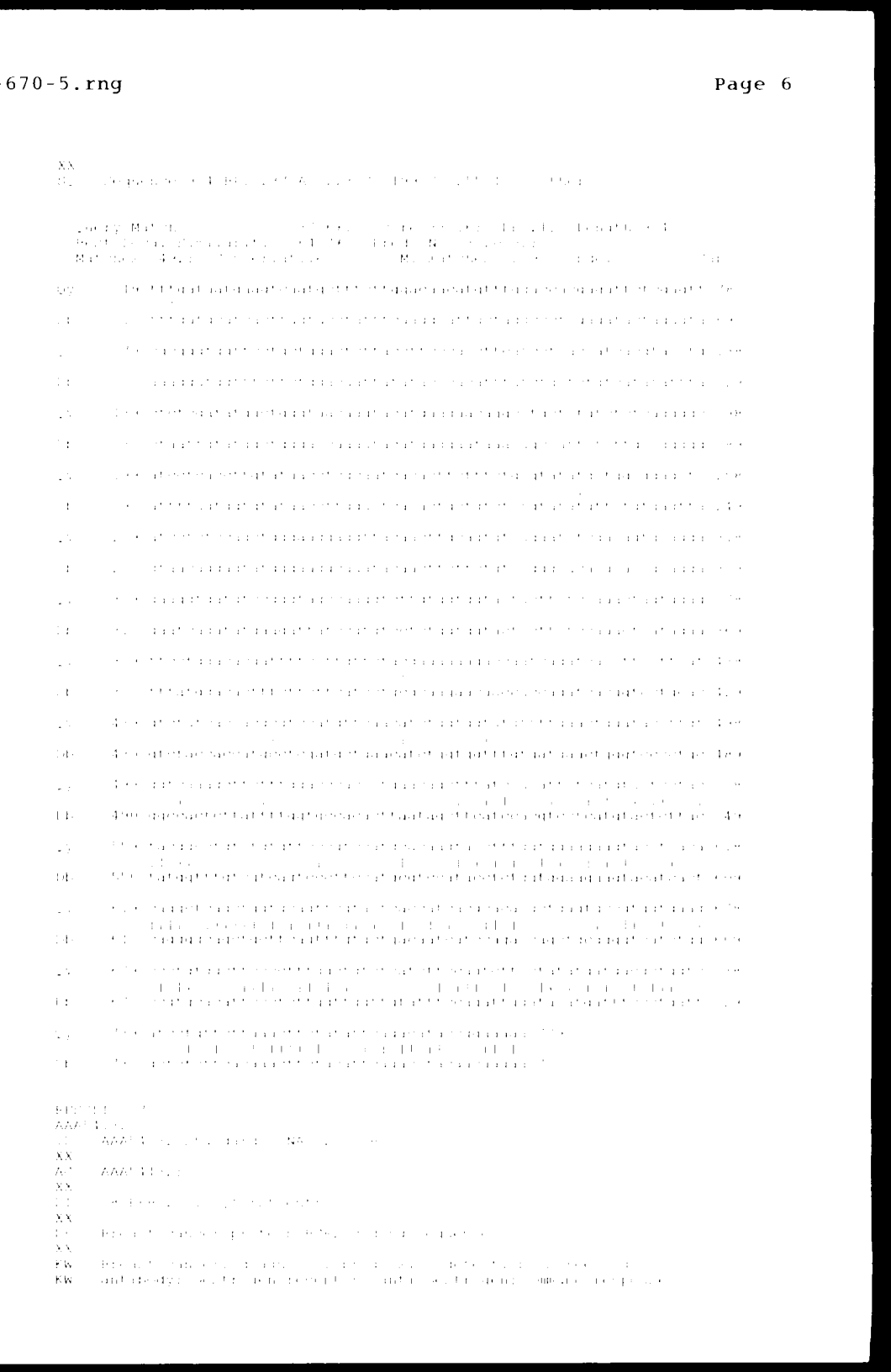




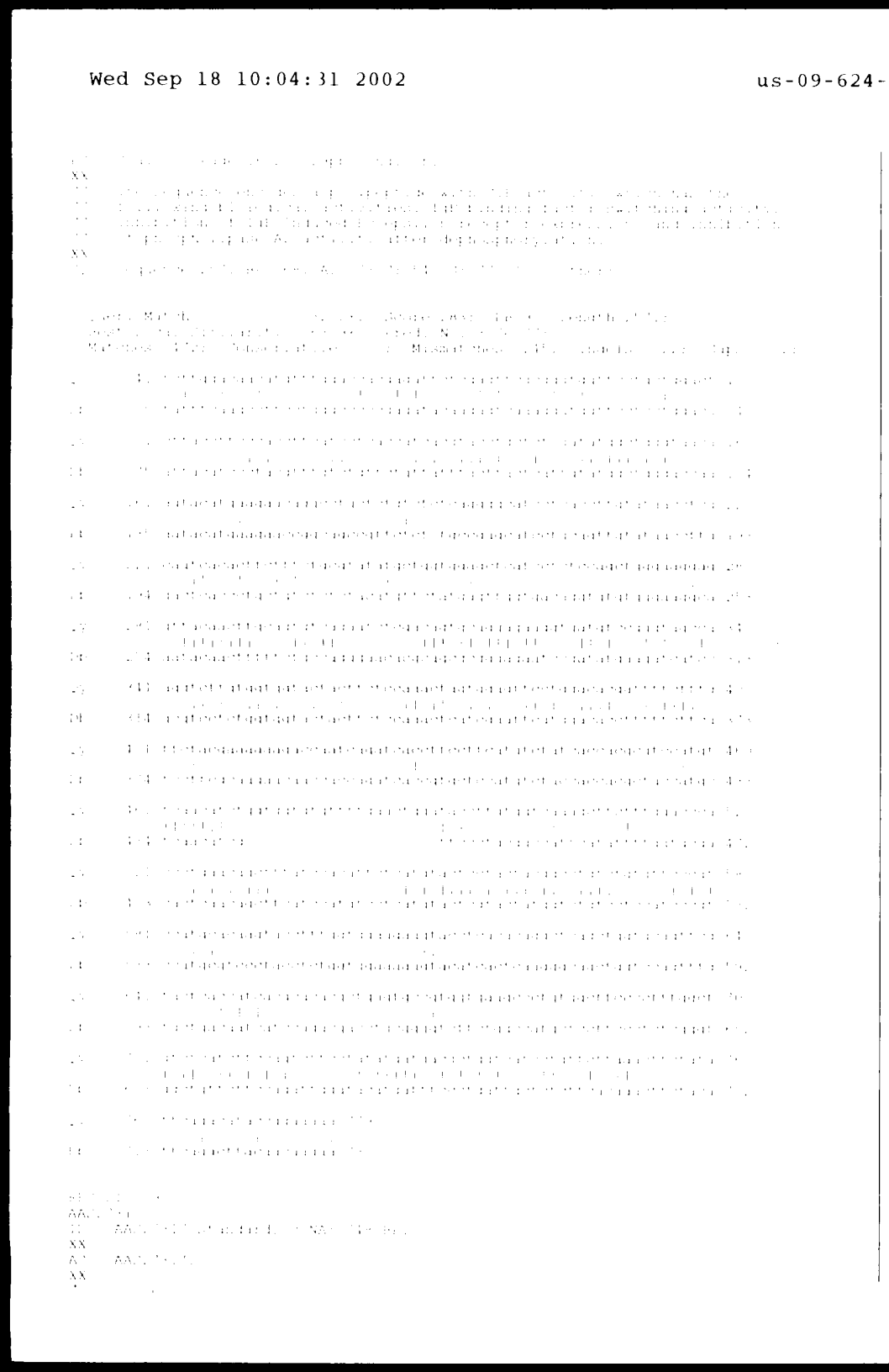


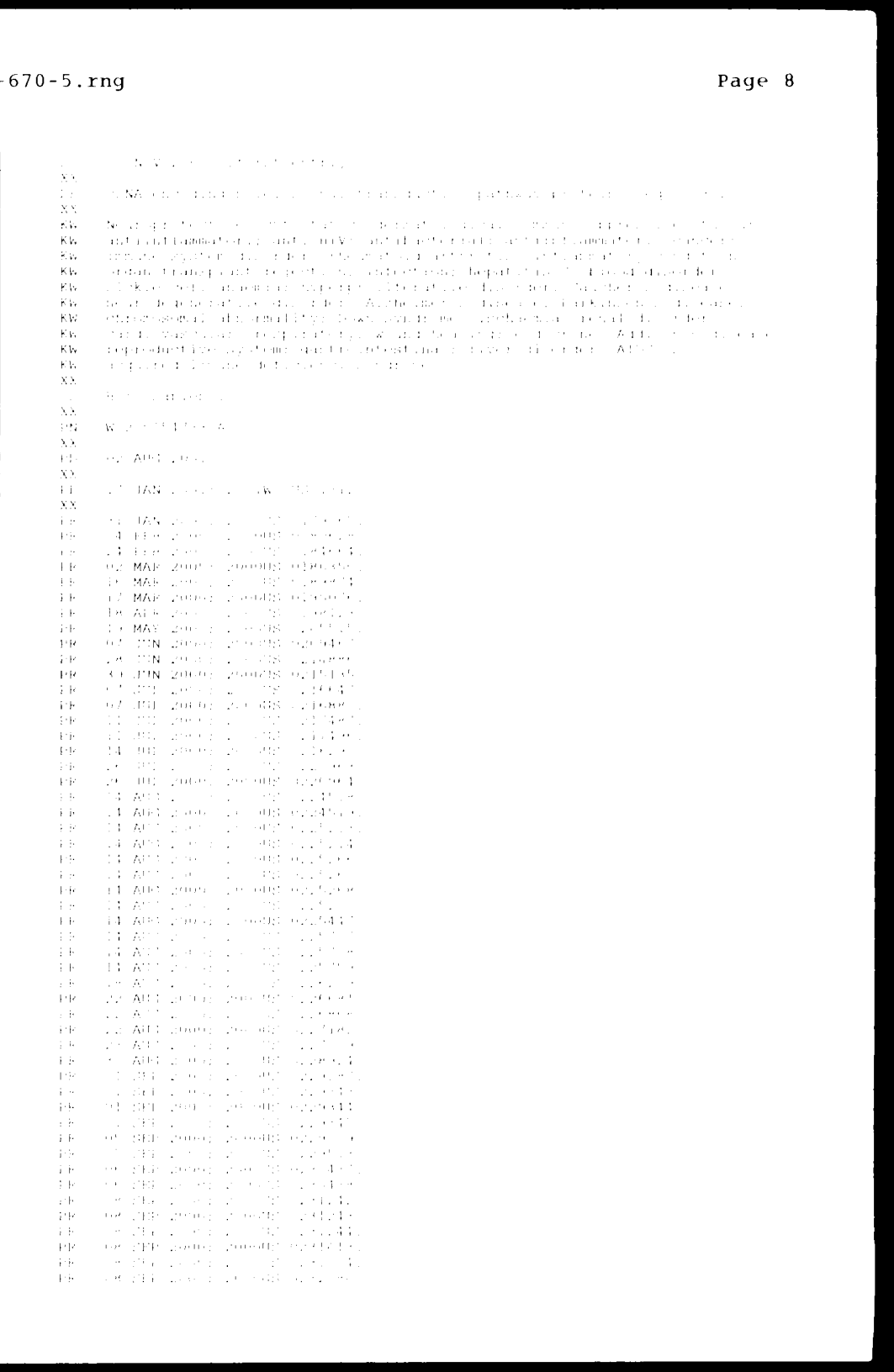






```
KW    [gene] knockout, multifunctional, receptor, B7R1, B7_1, B7_2, B7H1, B7N1, B7N2,
KW    B7N3, B7_1, B7X2, B7X3, B7A2, B7R2, B7_2, B7S1, human, ds.
XX
FT   Note: duplicate
XX
PN   Key                 Location/Qualifiers
FT   CDS                 1..1017
FT                       /function= "
FT                       /product= B7R1 protein
XX
PN   WO-A-9958658.
XX
PD   11-NOV-1999.
XX
PF   19-MAY-1999;  99WO-US010966.
XX
PR   19-MAY-1998;  98US-086078.
PR   12-NOV-1998;  98US-0108076.
PR   12-NOV-1998;  98US-0108077.
PR   20-NOV-1998;  98US-0109081.
PR   12-DEC-1998;  98US-0112047.
PR   19-MAY-2000;  2000US-0575187.
XX
PA   (BOSB ) BOS BIOTECHNOLOGY INC.
XX
PI   Mark L. Slish KC;
XX
DR   WPI; 2000-036216/01.
DR   P-PSDB; AAB95106.
XX
PT   Screening drug candidates for their ability to modulate breast cancer
PT   by contacting the drug to a cell expressing an expression profile gene
PT   and determining modulation of expression of the gene
XX
PS   Disclosure; Fig 6B; 258pp; English.
XX
CC   New methods for screening drug candidates are described which
CC   comprise adding a test candidate to a cell that expresses a protein
CC   selected from B7H1, B7A2, B7R1, B7N1, B7N2, B7S1, B7X2 and B7X3 or their fragments and determining the effect of the drug
CC   on the expression of these proteins. Antibodies to breast cancer
CC   genes (specifically B7H1 or its fragment (B7Hlp1 or B7Hlp2)) are
CC   useful for inhibiting and treating breast cancer in individuals who
CC   are non responsive to anti-oestrogen and positive for oestrogen
CC   receptor. Compositions comprising B7H1 or a nucleic acid encoding
CC   B7H1 are useful for eliciting an immune response in an individual.
CC   The antibodies are also useful for the diagnosis and prognosis of
CC   breast cancer and for screening compositions which modulate the
CC   breast cancer phenotype. The method allows rapid and simple
CC   detection of lymph node metastases.
XX
SQ   Sequence 271 BP; 81 A; 84 C; 52 G; 54 T; 0 other;


  Query Match             37.5%;  Score 629.6;  DB 2;  Length 271;
  Best Local Similarity   41.2%;  Pred. No. 1e-55;
  Matches  84;  Conservative   42;  Mismatches  27;  Indels   0;  Gaps   0;

Qy       1 tgtgtgaggaggcttgatttccaaggtggctttccatcaggtcgatgccagcctcagcgag 60
           |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 tgtgtgaggaggcttgatttccaaggtggctttccatcaggtcgatgccagcctcagcgag 60

Qy      61 ctcatcgtcaccaacagcgtgctgaacgagctcgtggacatccccgccctgcagctgtcg 120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ctcatcgtcaccaacagcgtgctgaacgagctcgtggacatccccgccctgcagctgtcg 120

Qy     121 gacgagtccaggctcttccaggacgcctacccccacgtcatctgggacgagtccgacggc 180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gacgagtccaggctcttccaggacgcctacccccacgtcatctgggacgagtccgacggc 180

Qy     181 atcctgcgcctgcagcagggctcccgcgccaccttcacctgcacggtgaagctgatcaag 240
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 atcctgcgcctgcagcagggctcccgcgccaccttcacctgcacggtgaagctgatcaag 240

Qy     241 ccccccgacgaggacggcgtgcccgtgcgctgcggcatccgccagtcctcagcccacctg 300
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ccccccgacgaggacggcgtgcccgtgcgctgcggcatccgccagtcctcagcccacctg 300

Qy     301 agcctgcagcccaacttcacctccgacgagcgcggcgagctgacctgcctggccgccgag 360
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 agcctgcagcccaacttcacctccgacgagcgcggcgagctgacctgcctggccgccgag 360

Qy     361 ctccacgtcgaccagctggagtcccccgtcaacctcacggtgcgctgcaacggcagcggc 420
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ctccacgtcgaccagctggagtcccccgtcaacctcacggtgcgctgcaacggcagcggc 420

Qy     421 gcctttcctgcgcccgaccgcgtcgtggagctggcccccctgggcaacgtgacgctgtcc 480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 gcctttcctgcgcccgaccgcgtcgtggagctggcccccctgggcaacgtgacgctgtcc 480

Qy     481 gtgcgcgcgcagcccgcgctgcgccacccgcgcatcatcgaggagatcgtgctgccgatc 540
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 gtgcgcgcgcagcccgcgctgcgccacccgcgcatcatcgaggagatcgtgctgccgatc 540

Qy     541 gcgctcgagcacgtgcgcgaggcgcaggtgcaggtgacggagcccgcgcagttcgccgcg 600
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 gcgctcgagcacgtgcgcgaggcgcaggtgcaggtgacggagcccgcgcagttcgccgcg 600

Qy     601 gccgtgcaggagcccagcgcgcagcgcacgatgtcctgggacctggacgtgcccttcgcc 660
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 gccgtgcaggagcccagcgcgcagcgcacgatgtcctgggacctggacgtgcccttcgcc 660

Qy     661 gccgtgcgcgcgcccagcgtgcagcgccgcgcgctgcagtgcgccctgcgcctgccgctg 720
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 gccgtgcgcgcgcccagcgtgcagcgccgcgcgctgcagtgcgccctgcgcctgccgctg 720

Qy     721 gcgcccgtgcgcgtgcgcctgcgcctggcc 751
           |||||||||||||||||||||||||||||
Db     721 gcgcccgtgcgcgtgcgcctgcgcctggcc 751


RESULT 1
AAN81168
ID   AAN81168 standard; DNA; 2571 BP.
XX
AC   AAN81168;
XX
DT   05-NOV-1999  (first entry)
XX
DE   Fragment from clone encoding polypeptide with glycosylation inhibiting
DE   factor activity.
XX
KW   Glycosylation; GIF; IGIF4; Th2; anti-inflammatory agent; ss.
XX
OS   Unidentified.
XX
RN   [1]
XX
RP   1-NOV-1999.
XX
RA   (ZYLI ) ZYMOGENETICS INC.
XX
RT   Ishizaka K, Mariano CL, Moore KW;
XX
DR   WPI; 1984-208165/40
DR   P-PSDB; AAB42500.
XX
CC   Nucleic acid encoding glycosylation inhibiting peptide capable of
CC   suppressing IgE immunoglobulin responses, for use as anti-inflammatory
CC   agents.
XX
```



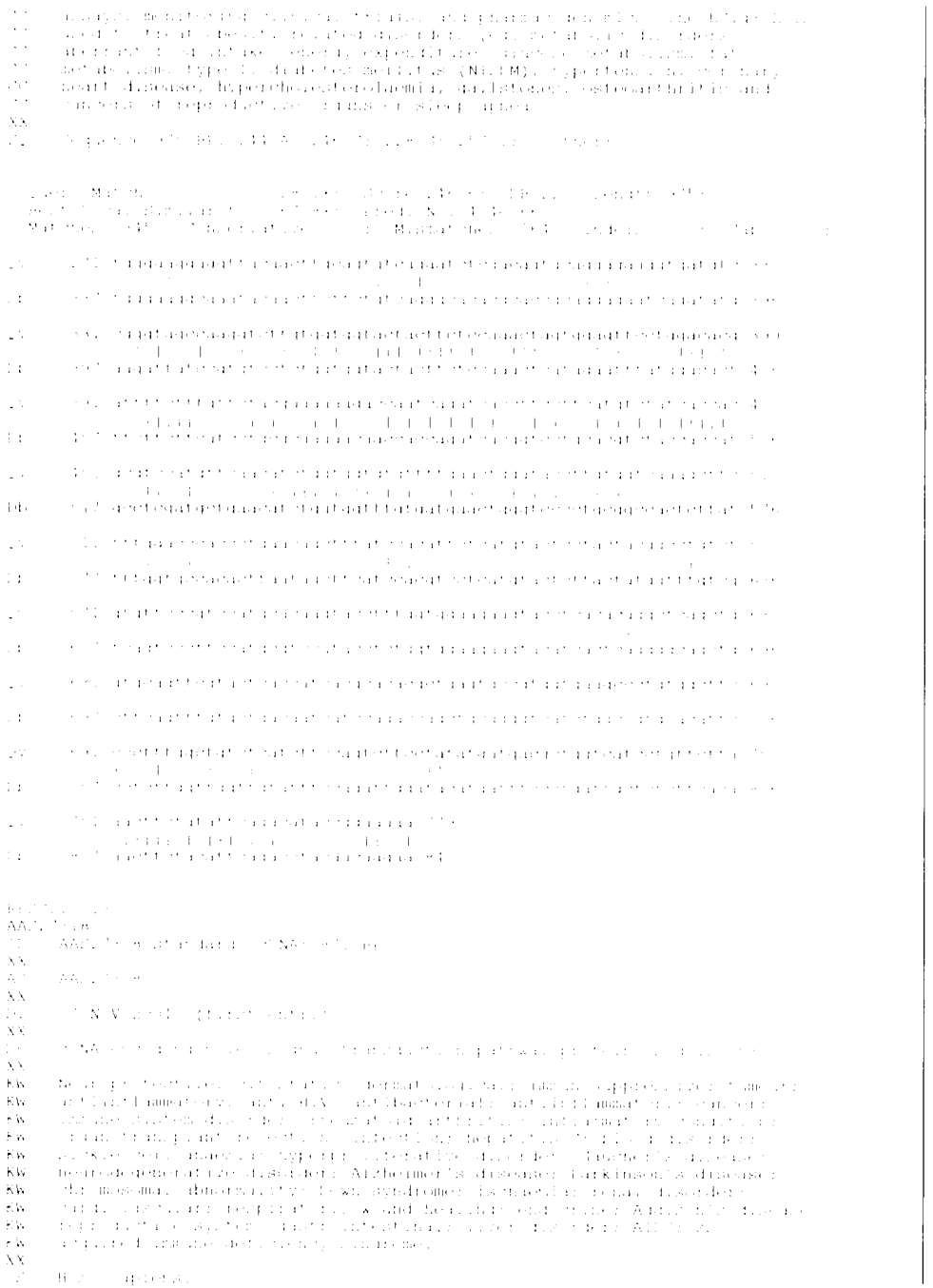


PN [illegible]
PD 12-SEP-2000; 2000US-0231968.
PD 14-SEP-2000; 2000US-0232547.
PD 14-SEP-2000; 2000US-0232548.
PD 14-SEP-2000; 2000US-0232899.
PD 14-SEP-2000; 2000US-0232900.
PD 14-SEP-2000; 2000US-0232901.
PD 14-SEP-2000; 2000US-0232903.
PD 14-SEP-2000; 2000US-0232904.
PD 14-SEP-2000; 2000US-0233065.
PD 21-SEP-2000; 2000US-0234223.
PD 21-SEP-2000; 2000US-0234274.
PD 25-SEP-2000; 2000US-0234907.
PD 25-SEP-2000; 2000US-0234908.
PD 26-SEP-2000; 2000US-0235484.
PD 27-SEP-2000; 2000US-0235834.
PD 27-SEP-2000; 2000US-0235835.
PD 28-SEP-2000; 2000US-0236527.
PD 28-SEP-2000; 2000US-0236607.
PD 28-SEP-2000; 2000US-0236688.
PD 29-SEP-2000; 2000US-0236869.
PD 29-SEP-2000; 2000US-0236875.
PD 29-SEP-2000; 2000US-0236882.
PD 29-SEP-2000; 2000US-0237037.
PD 29-SEP-2000; 2000US-0237038.
PD 29-SEP-2000; 2000US-0237059.
PD 29-SEP-2000; 2000US-0237341.
PD 29-SEP-2000; 2000US-0237665.
PD 29-SEP-2000; 2000US-0237667.
PD 29-SEP-2000; 2000US-0241221.
PD 29-SEP-2000; 2000US-0241795.
PD 29-SEP-2000; 2000US-0241796.
PD 29-SEP-2000; 2000US-0241797.
PD 29-SEP-2000; 2000US-0241798.
PD 29-SEP-2000; 2000US-0241799.
PD 29-SEP-2000; 2000US-0241802.
PD 01-NOV-2000; 2000US-0244127.
PD 06-NOV-2000; 2000US-0246474.
PD 06-NOV-2000; 2000US-0246475.
PD 06-NOV-2000; 2000US-0246476.
PD 06-NOV-2000; 2000US-0246477.
PD 06-NOV-2000; 2000US-0246479.
PD 06-NOV-2000; 2000US-0246523.
PD 08-NOV-2000; 2000US-0246624.
PD 08-NOV-2000; 2000US-0246625.
PD 08-NOV-2000; 2000US-0246626.
PD 08-NOV-2000; 2000US-0246627.
PD 08-NOV-2000; 2000US-0246628.
PD 08-NOV-2000; 2000US-0246642.
PD 08-NOV-2000; 2000US-0246679.
PD 08-NOV-2000; 2000US-0246711.
PD 08-NOV-2000; 2000US-0246713.
PD 08-NOV-2000; 2000US-0246714.
PD 13-NOV-2000; 2000US-0247207.
PD 13-NOV-2000; 2000US-0247208.
PD 13-NOV-2000; 2000US-0247214.
PD 13-NOV-2000; 2000US-0247215.
PD 13-NOV-2000; 2000US-0247216.
PD 13-NOV-2000; 2000US-0247217.
PD 13-NOV-2000; 2000US-0247218.
PD 13-NOV-2000; 2000US-0247219.
PD 13-NOV-2000; 2000US-0247220.
PD 13-NOV-2000; 2000US-0247221.
PD 13-NOV-2000; 2000US-0247222.
PD 13-NOV-2000; 2000US-0247223.
PD 13-NOV-2000; 2000US-0247224.
PD 13-NOV-2000; 2000US-0247225.
PD 13-NOV-2000; 2000US-0247226.
PD 13-NOV-2000; 2000US-0247227.
PD 13-NOV-2000; 2000US-0247228.
PD 13-NOV-2000; 2000US-0247229.
PD 13-NOV-2000; 2000US-0247230.
PD 13-NOV-2000; 2000US-0247231.
PD 13-NOV-2000; 2000US-0247232.
PD 13-NOV-2000; 2000US-0247233.
PD 13-NOV-2000; 2000US-0247234.
PD 13-NOV-2000; 2000US-0247235.
PD 13-NOV-2000; 2000US-0247236.
PD 13-NOV-2000; 2000US-0247237.
PD 13-NOV-2000; 2000US-0247238.
PD 13-NOV-2000; 2000US-0247239.
PD 13-NOV-2000; 2000US-0247240.
PD 13-NOV-2000; 2000US-0247241.
PD 13-NOV-2000; 2000US-0247242.
PD 13-NOV-2000; 2000US-0247243.
PD 13-NOV-2000; 2000US-0247244.
PD 13-NOV-2000; 2000US-0247245.
PD 13-NOV-2000; 2000US-0249264.
PD 13-NOV-2000; 2000US-0249265.
PD 13-NOV-2000; 2000US-0249267.
PD 13-NOV-2000; 2000US-0249269.
PD 13-NOV-2000; 2000US-0249270.
PD 04-DEC-2000; 2000US-0250761.
PD 04-DEC-2000; 2000US-0250762.
PD 05-DEC-2000; 2000US-0251040.
PD 05-DEC-2000; 2000US-0251088.
PD 05-DEC-2000; 2000US-0256719.
PD 06-DEC-2000; 2000US-0251479.
PD 07-DEC-2000; 2000US-0251866.
PD 08-DEC-2000; 2000US-0251868.
PD 08-DEC-2000; 2000US-0251869.
PD 08-DEC-2000; 2000US-0251984.
PD 08-DEC-2000; 2000US-0251990.
PD 11-DEC-2000; 2000US-0254047.
PD 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA) HUMAN GENOME SCI INC.
XX
IC Ensor CA; Barash SC; Ebner SM;
XX
PR WPI; 2001-465491/50.
PR PS26; AA12740 .
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders
XX
PS Claim 1; SEQ ID NO 552; 869pp; English.
XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful in
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. severe combined immunodeficiencies, i.e. SCIDs,
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejection, graft versus host disease, infectious diseases
CC (e.g. hepatitis), bleeding disorders, haemoglobin abnormalities and
CC other blood related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, to
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AA12740-AA12750 represent novel signal transduction
CC pathway protein coding sequences and PCR primers of the invention.
XX

Score 2708; DB 22; Length 7496;
Best Local Similarity 64.0%; Pred. No. 1.1e-60;
Matches 982; Conservative 1; Mismatches 217; Indels 0; Gaps 0;

```
RF   ...-JUL-2000; 2000JP-0214069.
XX
PR   ...-JUL-1999;   99JP-0194486.
PR   11-JAN-2000; 2000JP-0018774.
PR   ...-MAY-2000; 2000JP-0161766.
XX
PA   (HELI ) HELIX RES INST.
XX
PI   Ito I,  Nishikawa T,  Hosoki T,  Hayashi K,  Ishii S,  Kawai Y;
PI   Wakamatsu A,  Otsuki T,  Nagai K,  Kimura S,  Ohsaki E,  Koga H;
XX
DR   WPI; 2001-524255/58.
XX
PT   New primers useful for synthesizing full length cDNA clones and their
PT   use in genetic manipulation.
XX
PS   Example 11; SEQ ID NO 2409; 1380pp + sequence listing; English.
XX
CC   The invention relates to primers for synthesising full length cDNA
CC   clones. New cDNA molecules encoding a human protein have been
CC   isolated and nucleotide sequences of 5' and 3' ends of the cDNA
CC   molecules have been determined. Primers for synthesising the full length
CC   cDNA are useful for clarifying the function of the protein encoded by
CC   the cDNA. The full length clones were obtained by construction of full
CC   length enriched cDNA libraries that were synthesised by the oligo capping
CC   method. The primers enable the production of the full length cDNA easily
CC   without any special methods. The present sequence was used as the
CC   representative sequence from a human clone which was used in
CC   homology searches to identify the clone.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ   Sequence 680 BP; 167 A; 164 C; 144 G; 149 T; 6 other;
```

```
  Query Match           28.8%;  Score 254;  DB 22;  Length 680;
  Best Local Similarity   64.8%;  Pred. No. 3.7e-68;
  Matches  419;  Conservative    0;  Mismatches  224;  Indels    4;  Gaps    3;

Qy        7 cagctgaaggcctttgataatgaaatcaatactttctgaaacaacatgttgaaccaaga  66
            |.  |:      ||||||    | | | | | | | ||  |     || | ||| || |||
Db        3 caaatgaaacattttgatgcatcactagtacctatttcaaggacattgctgtaggcccaga  62

Qy       67 gattctcgaattcaagcagtgatttcctgctgaacatcttaccttcccaccttcatcctcacc  126
            ||: || ||||    ||:||||| || |||||  ||| | ||||| || ||| | | |
Db       63 gatactaggagtaaaaaaatgatttcttcttcagcaattatatacccacattgatctgctct  122

Qy      127 atcacgtaccctgctctcgatatgctgttaacaagtacatgaaaaacagagctgtctct  186
            :|| || | ,|     ||||||||||    ||:|||||| ||| | |   ||| |
Db      123 gtcatatattactaattatatgcgaaaaaaaaaatgtgtgaaaatcaaaaaaaaaaaaaa  182

Qy      187 tctctcagaggcatctctgcttatgaacctgcaaataacctcttttctgaaatatg  246
            ||| | |||| || | |  ||||:|||||| | | ||||||| || |||  |||||||
Db      183 tctgcaaggaaatttaatgatgtgataacattgaatcaaatagtgtctgtatatata  242

Qy      247 gtgatgaactcatcctctgaagctgaggaagaatgttacaacttgcaatgtagaatctc  306
            |   |||| |           || |||||   |||||||   ||||||
Db      243 ttctgtgaattaatgaacaggatgtgggaaaacaaatacaacttcttctgtcaggcaca  302

Qy      307 gacagtgcaggaaaaaatgtgatgtcccagtaagccaaaatctttatagtgatgatactacttctcc  366
            ||  ||||||||  ||| | | |  |  ||| | ||||||||||||||||||
Db      303 cgtacgcagaggaaaaatcaatatgaaaattatccgtatcctctgatgatactacttctcc  362

Qy      367 aaactatgaattcctcagacaccatttcttatctatagaaaaaaacaatcaaatc  426
            ||||    || |  ||:|||  | ||||| | || ||| || | | | ||||||
Db      363 aaactctatgaatttatgaacacttctcttctatcctcacaaaaacaaccacctaatc  422

Qy      427 accttcctcgatgtatcaccacgtccatgttcaaatctagtgatgtatgtttgaac  486
            |:  ||| | ||||| |||| || | | ||| | ||||||||||||| |||  |||||
Db      423 acagtcctcgacatctaccacgatgcctcgatgctgaacatctgatgattgtgatgaac  482

Qy      487 tgagtaccttgtgatcaaaagttctttgaccacctgaacagtttatccaattctc  546
            ||| | || || ||:||    | ||||| |||| || || ||||| |||||| | |||
Db      483 tgagtcccctgcggcactcttatttgatgccaccttaataagttcatccaagtcctc  542

Qy      547 atgtactcctactacagcctatctgt-gttcccgtccatgcacaatacctttgatgg  605
            |||||||| ||||| || |||| ||     ||| |||||||     |||||  ||||
Db      543 atgtactcttactatgagtttgtccatcagtccctttcatgcgtccataccttcagtgaaa  602

Qy      606 gatgaatgcctcgacagcaggctcatctgtgcagttctactcacca  644
            ||||||||| |'|| | ||  |. ||| | || || | |  || |
Db      603 aagaaatgcctcactcaggtgcacatgctgctcatttgtacagcaa  649


RESULT  12
AAB07290
ID   AAB07290 standard; cDNA; 975 BP.
XX
AC   AAB07290;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Human cold-induced glycoprotein 30-like protein (hCig30L) cDNA.
XX
KW   Human; cold-induced glycoprotein 30-like protein; hCig30L; anorectic;
KW   drug screening; gene therapy; restriction fragment length polymorphism;
KW   RFLP; chromosomal mapping; tissue typing; forensic; metabolic disorder;
KW   obesity related disorder; aberrant food intake; energy expenditure;
KW   glucose metabolism; fat metabolism; type II diabetes mellitus; NIDDM;
KW   hypertension; coronary heart disease; hypercholesterolaemia; gallstone;
KW   osteoarthritis; cancer; sleep apnea; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..975
FT                   /*tag= a
FT                   /product= "Human cold-induced glycoprotein 30-like
FT                   protein (hCig30L)"
XX
PN   WO200131013-A1.
XX
PD   03-MAY-2001
XX
PF   27-OCT-2000; 2000WO-US29559.
XX
PR   28-OCT-1999;  99US-0161978.
PR   26-OCT-2000;  2000US-3333444.
XX
PA   (CORA ) CURAGEN CORP.
XX
PI   Lewin DA,  Stewart TA;
XX
DR   WPI; 2001-328652/34.
DR   P-PSDB; AAE02842.
XX
PT   Human cold-induced glycoprotein like-30 protein and nucleic acids
PT   encoding the protein useful for treating obesity-related disorders,
PT   e.g. metabolic disorders, aberrant food intake, energy expenditure or
PT   fat metabolism -
XX
PS   Claim 8; Fig 2A; 97pp; English.
XX
CC   The present sequence is human cold induced glycoprotein 30-like protein
CC   (hCig30L) cDNA. The hCig30L proteins are useful as immunogens to raise
CC   anti hCig30L antibodies, to screen drugs or compounds that modulate
CC   protein activity or expression and to treat disorders characterised by
CC   insufficient or excessive production of an hCig30L protein or production
CC   of hCig30L protein forms that have decreased or aberrant activity
CC   compared to wild-type protein. The nucleic acids are also used to
CC   express hCig30L protein for gene therapy applications, to detect hCig30L
CC   mRNA in a biological sample or a genetic lesion in hCig30L and as
CC   additional DNA markers for restriction fragment length polymorphisms
CC   (RFLP). The hCig30L is used in screening assays, in detection assays
CC   such as chromosomal mapping, cell and tissue typing, forensic and
CC   biology and in preventive medicine (e.g. diagnostic or prognostic
```



PR 27-SEP-2000; 2000US-0241221.
PR 27-SEP-2000; 2000US-0241785.
PR 27-SEP-2000; 2000US-0241786.
PR 27-SEP-2000; 2000US-0241787.
PR 27-SEP-2000; 2000US-0241808.
PR 27-SEP-2000; 2000US-0241809.
PR 27-SEP-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244117.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246479.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249269.
PR 17-NOV-2000; 2000US-0249270.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250341.
PR 05-DEC-2000; 2000US-0251698.
PR 05-DEC-2000; 2000US-0251699.
PR 05-DEC-2000; 2000US-0251700.
PR 06-DEC-2000; 2000US-0251847.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251899.
PR 08-DEC-2000; 2000US-0251900.
PR 07-DEC-2000; 2000US-0254107.
PR 29-JAN-2001; 2001US-0265078.

XX

PA (HUMA) HUMAN GENOME SCI INC.

XX
XX
PI Rosen A, Barash S, Bates SM;
XX

DR WPI; 2001-466460/50.
DR P-PSDB; AAC74401.

XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing diseases related to the proteins, including cancers, immune
PT disorders and neuronal disorders
XX
PS Claim 1; SEQ ID NO 35; 86pp; English.
XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II), (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheuma oid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (Aids). AAS27697 AAS27650 represent novel signal transduction
CC pathway protein coding sequences and PCR primers of the invention.
XX

; Score: 67.44; Score 241; DB 22; Length 807;
Best Local Similarity: 67.8%; Pred. No. 2.8e-64;
Matches 847; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

Qy 283 tacaaatgcaatatgtcaggaatctcaacaatgcaaggaaggaatgaatgtcaggtaacaaaa 342
||||||| |||||| || |||||||| || | || | |
Db 2 tacaaatccttctgtcaggggcacagcacgcgcaaggaatgcaatacaaggattccgt 60

Qy 343 gtcttgtatgtatactacttctcaaacaatgcgattctgcgacagtatttttgtttt 402
|| | |.|.| |||| |||| |||||||| | || || |||||| | ||||| |
Db 61 gtccttgtgtgtgatacttctccaaacctgtaggattatgcgacacttcttcttcatc 120

Qy 403 ctacaaaaaaaaacaacaatgcaatgcaccttccttcatgtctatcaacaagcgtccaatttc 462
|| || || | | | ||||||| |||| || ||||| |||| || || ||
Db 121 ctgcgcaaaacaaccaaatgcagtcctgcaccatgcaacatgccgtccaatta 180

Qy 463 aacctctgtgtgtgtttttaaactgaatacctttgtgttcaaagcttctttgaccaccc 522
|||| |||||||| || | | | | | | | || || | | |||| | |
Db 181 aacttctgtgtgtttttgtgaaaacgatgtccttgtgcgcactcttatttttgatgccaca 240

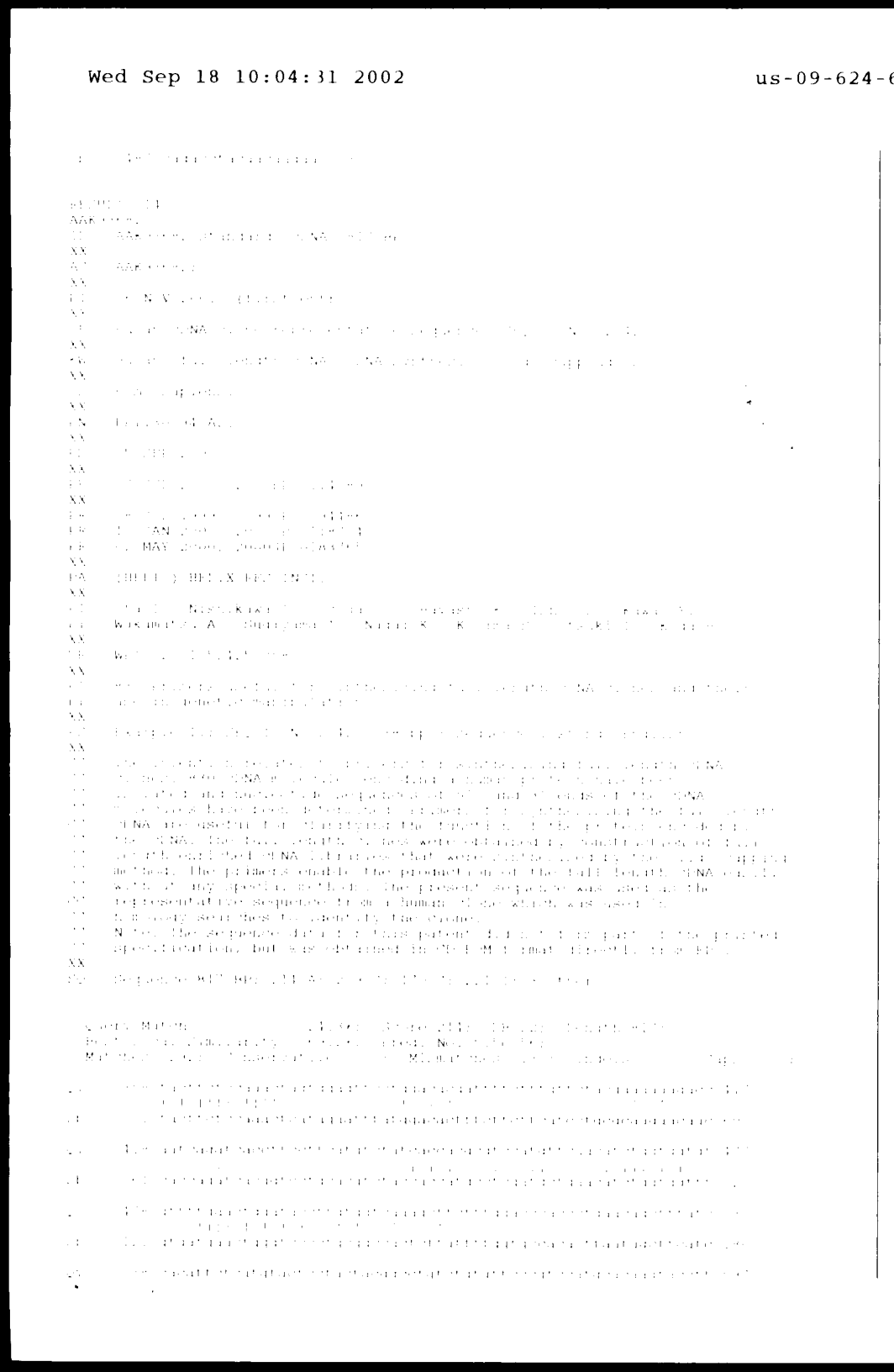
Qy 523 ctaaccctttatcccaattctcatgtatgtctactacgccgtctatgtcgttcccgtcc 582
|| | | | ||||| | ||||||| || |||| || |||| |||| ||
Db 241 ctttatgcctttcaccgtctcctatgtactcttctatgattgtatgtcgatccctccgtc 300

Qy 583 atgcgcaaatacttgttgtgaaaaaatacctcacgcaggctcgctgtgtcgattcctata 642
|||| |||| ||||||||||||| |||| |||| |||||| | |||| |
Db 301 atacctcaatacctctgtgtgaaaaatatcatcatcactcgggcgctacttcgattttata 360

Qy 643 ctcaccatacgcgacgcgtgatgcgtgtgcaaaccttgtgcttcctttgcgtat 702
|| | ,|| | || | | || | | || || |||| |||| ||
Db 361 ctgtccatcatgcgtcaaccaacgtgcaagtcatcttgcgtgtgcattccttcttgctttca 420

Qy 703 ctcatctcgatcgatcgtatatatatcactgatccatcctgttcttaaacttctatttt 762
| ||||| | || || | | | | |||||||| |||
Db 421 ttgtatttcgaatgcgatgcatcgatttcctgattgtctct cgcaaacttgtacat 480

Qy 763 caaatatcccaaaaaa 779
|||| || || ||



CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1203 BP; 287 A; 294 C; 290 G; 332 T; 0 other;

Query Match 20.2%; Score 177.6; DB 23; Length 1203;
Best Local Similarity 55.1%; Prob. No. 1.4e-44;
Matches 392; Conservative 0; Mismatches 314; Indels 6; Gaps 2;

Qy 86 aaattctcaaattcaaaaatgattcctgtgaaattacctcccacctcatcctcac 135
Db 857 aaataaacatatgaaaaattgacctctgatgcaattctcttgacctacactaaatataaa 916

Qy 136 catcacatacctactctcaatatgactgaataacaaatacatgaaaaacaaactgctct 185
Db 917 cattctttatctcctgtttatatgtgctgaaaccaaaatgaatgaaaaaccagaaaccttt 976

Qy 186 gtctctcaaagacatcctcaccttatataacctcacaatcacacttctttctgcatatat 245
Db 977 tcagatgctctatgctcattatctataatttgaagatgattttgctttaacctcttat 1036

Qy 246 gctgttgaagctcatcctccttccaactgaaaaaaaattacaacttgcagtatcaaaatct 305
Db 1037 cttcaaaagttatttcatgaatatatgatgcagaataagctatatttgccagagtgt 1096

Qy 306 cgacaatgcaaaaaaaatgat---gtccagtagccaaagatctttgtgttgtatactactt 362
Db 1097 gaatttattctaataatattcatgaaatcaagatagctgctgctctgtggtgtacttgt 1156

Qy 363 ctccaaactagtgaatttcctgaacacgatttcttttgttctacgaaaaaaqaccaatca 422
Db 1157 atctaaaaaagttgaatatttgaacacaatgtttttattctgaaaaaaaaaacaacca 1216

Qy 423 gatcacctccttcatgtctatcaccacgcgtcatgttcaaaatctgatgtgtgtgtttt 482
Db 1217 aatttcttcctcatgtgtatcatcactgtacgatgttacctttgtgatgatgaaat 1276

Qy 483 gaactgaatacctttatgatcaaagcttctttgtacccaccctgaacagctttatccacat 542
Db 1277 taattgagttgcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1336

Qy 543 tctcatatatcctctactacagcgctgtctatgtt---cccatccatcacaagtacctttg 599
Db 1337 gtatatatgactcatactatgagattaactgcattgagcccatgattcaaaaatatctttg 1396

Qy 600 gtgaaaaagtacctcacaagactcagctgtatcagttcatactcaccatcacgcacac 659
Db 1397 gtgaaaacatacctgaatatgttgcaactgattcaattccatgtgaccattgaacacac 1456

Qy 660 gctgagtacgtatgaaacccatgacttcccctttagctgtctcatctccaattctc 719
Db 1457 gacactgtctctttcacactcatgacccctcccaaatgaatgctgctgaagtctaattgc 1516

Qy 720 ?? 771
Db 1517 ??? 1068

Search completed: September 17, 2002, 18:47:28
Job time: 5617 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: September 17, 2002, 18:01:43 ; Search time 70.12 Seconds
(without alignments)
3152.736 Million cell upda.os/sec

Title: US-09-624-670-6
Perfect score: 900
Sequence: 1 atgggaacattcgatgatgc..... ..aaaaaaaacaaaaaaaaatttgaa 900

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ispna/PCTUS_Rnew.seq:*
2: /cgn2_6/ptodata/1/ispna/PCTUS_Rnew1.seq:*
3: /cgn2_6/ptodata/1/ispna/US06_COMB.seq:*
4: /cgn2_6/ptodata/1/ispna/US07_COMB.seq:*
5: /cgn2_6/ptodata/1/ispna/US08_COMB.seq:*
6: /cgn2_6/ptodata/1/ispna/US09_Rnew.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	98.4	10.9	124	4	US-09-172-108-45	Sequence 45, Appl
2	62.6	7.0	989	1	US-07-885-970A-7	Sequence 7, Appli
3	62.6	7.0	989	1	US-08-03 298 687A-7	Sequence 7, Appli
4	62.6	7.0	989	1	US-08-530-797-6	Sequence 6, Appli
5	62.6	7.0	989	1	US-08-298-829-7	Sequence 7, Appli
6	62.6	7.0	989	2	US-08-787-335-6	Sequence 6, Appli
7	55.8	6.2	7213	1	US-08-08 232-462-14	Sequence 14, Appl
8	43.2	4.8	289	4	US-09-007-005-17	Sequence 17, Appl
9	43.2	4.8	289	4	US-09-244-796-17	Sequence 17, Appl
10	41.2	4.6	2277	1	US-08-676-967-2	Sequence 2, Appli
11	41.2	4.6	2277	1	US-08-676-974-2	Sequence 2, Appli
12	41.2	4.6	2277	2	US-09-098-487-2	Sequence 2, Appli
13	38.8	4.3	1431	6	5245018-5	Patent No. 5245018
14	38.8	4.3	1801	1	US-07-847-562-1	Sequence 1, Appli
15	38.8	4.3	1801	1	US-08-240-328-1	Sequence 1, Appli
16	38.8	4.3	1801	2	US-08-215-089-8	Sequence 8, Appli
17	38.8	4.3	1801	2	US-08-990-849-1	Sequence 1, Appli
18	38.8	4.3	1801	5	PCT-US95-03384-8	Sequence 8, Appli
19	37.4	4.2	2493	1	US-07-977-434-5	Sequence 5, Appli
20	37.4	4.2	2493	1	US-08-458-819-5	Sequence 5, Appli
21	37.4	4.2	2493	5	PCT-US91-07035-5	Sequence 5, Appli
22	37.2	4.1	620	2	US-08-215-089-1	Sequence 1, Appli
23	37.2	4.1	620	5	PCT-US95-03384-1	Sequence 1, Appli
24	37.2	4.1	1600	1	US-08-073-384-12	Sequence 12, Appl
25	37.2	4.1	1600	1	US-08-254-359A-12	Sequence 12, Appl
26	37.2	4.1	1600	1	US-08-483-043-12	Sequence 12, Appl
27	37.2	4.1	1600	1	US-08-461-256-12	Sequence 12, Appl
28	37.2	4.1	1600	2	US-08-471-066A-12	Sequence 12, Appl
29	37.2	4.1	1600	2	US-08-484-956-12	Sequence 12, Appl
30	37.2	4.1	1600	2	US-08-757-653-12	Sequence 12, Appl
31	37.2	4.1	1600	2	US-08-549-491-12	Sequence 12, Appl
32	37.2	4.1	1600	2	US-08-756-486-12	Sequence 12, Appl
33	37.2	4.1	1600	2	US-08-823-516-12	Sequence 12, Appl
34	37.2	4.1	1600	3	US-08-682-853A-12	Sequence 12, Appl
35	37.2	4.1	1600	3	US-08-759-038-12	Sequence 12, Appl
36	37.2	4.1	1600	3	US-08-758-314-12	Sequence 12, Appl
37	37.2	4.1	1600	4	US-09-350-309-12	Sequence 12, Appl
38	37.2	4.1	1682	3	US-09-096-399-1	Sequence 1, Appli
39	37.2	4.1	1686	2	US-08-648-657-1	Sequence 1, Appli
40	37.2	4.1	1689	2	US-08-648-657-2	Sequence 2, Appli
41	37.2	4.1	1696	3	US-09-096-399-3	Sequence 3, Appli
42	37.2	4.1	2499	1	US-07-977-434-1	Sequence 1, Appli
43	37.2	4.1	2499	1	US-08-458-819-1	Sequence 1, Appli
44	37.2	4.1	2499	3	US-07-602-848E-1	Sequence 1, Appli
45	37.2	4.1	2499	4	US-09-587-856-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-172-108-45
; Sequence 45, Application US/09172108
; Patent No. 6160104
; GENERAL INFORMATION:
; APPLICANT: Cunningham, Mary Jane
; APPLICANT: Zweiger, Gary R.
; APPLICANT: Panzer, Scot R.
; APPLICANT: Seilhamer, Jeffrey J
; TITLE OF INVENTION: MARKERS FOR PERIOXISOMAL PROLIFERATORS
; FILE REFERENCE: DA 0012 US
; CURRENT APPLICATION NUMBER: US/09/172,108
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PERL Program
; SEQ ID NO 45
; LENGTH: 124
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 84, 93, 87, 109, 122
; OTHER INFORMATION: a or g or t, unknown, or other
; FEATURE:
; OTHER INFORMATION: 700535905H1
US-09-172-108-45

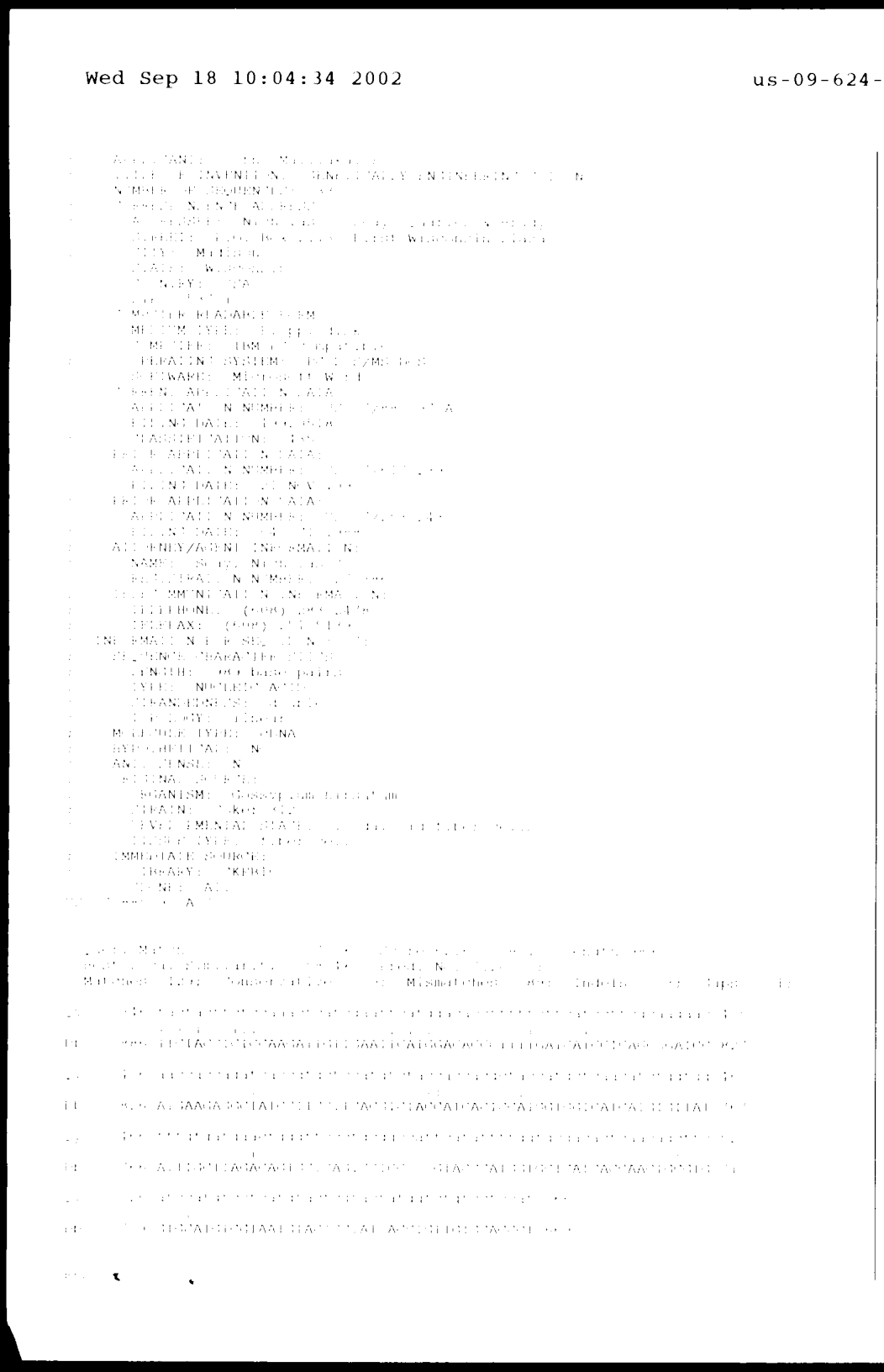
Query Match 10.9%; Score 98.4; DB 4; Length 124;
Best Local Similarity 91.1%; Pred. No. 1.3e-21;
Matches 112; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

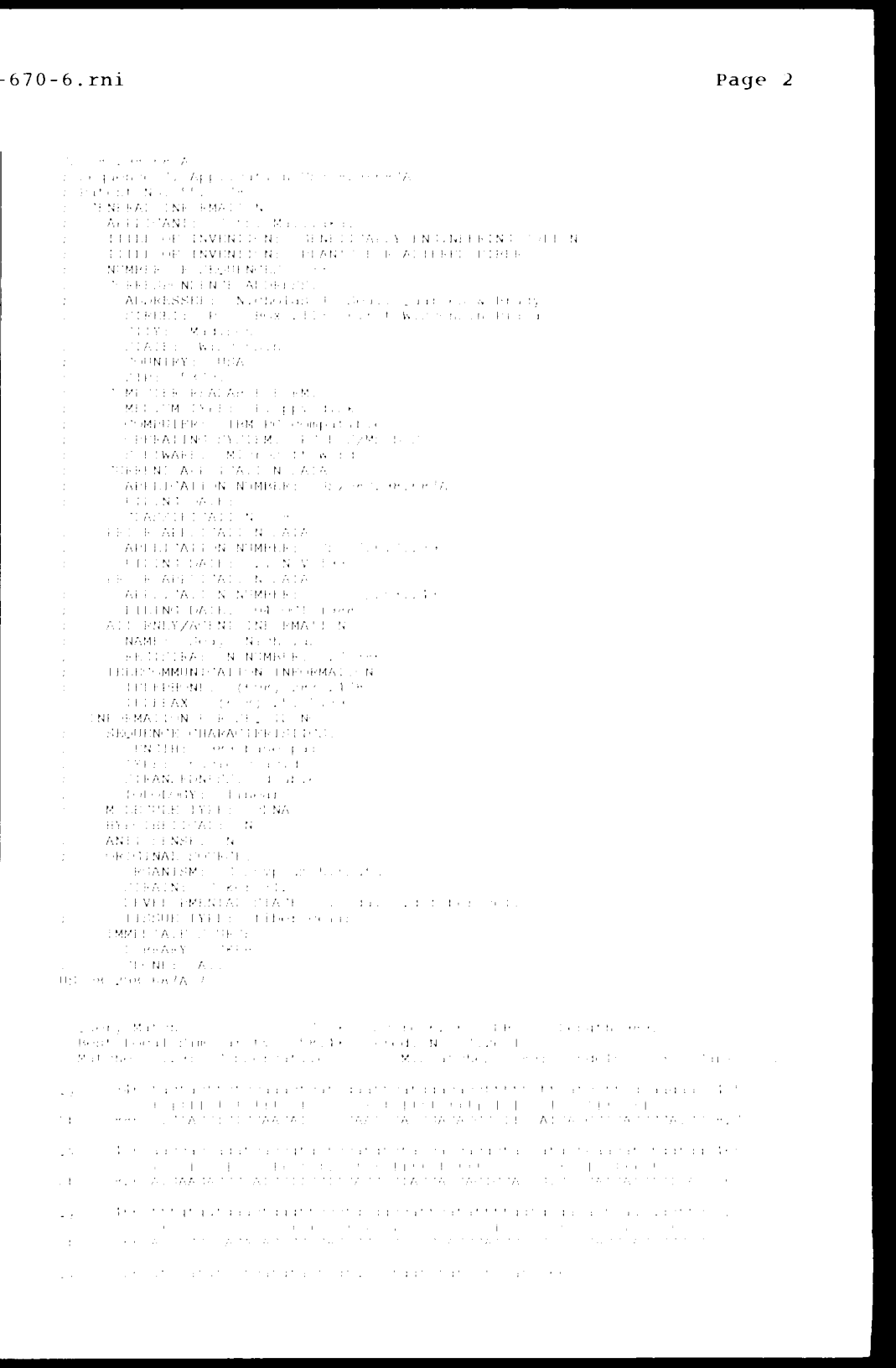
Qy 581 cctacctctgtgtgaaaaaatactcatcactcaaggcgcagctggtccaatttgt-gctgaca 639
||||||||||||||||||||||||||||| | |||||||||||||||||| |||||||
Db 1 cctacctctgtgtgaaaaaatactcatcactcaaggcgcagctggtccaatttgtagctgaca 60

Qy 640 atcctcaaacaactctcaagtatctctgcccatgctcctcccctctcgggtgactattc 699
|||||||||| | |||||||||| | ||||| |||||||||||||||| |||||||| |
Db 61 atcatccaaacaagctgcgggtcatatgacctgctcctccctctcaatgactgtac 120

Qy 700 ttcc 703
| ||
Db 121 tacc 124

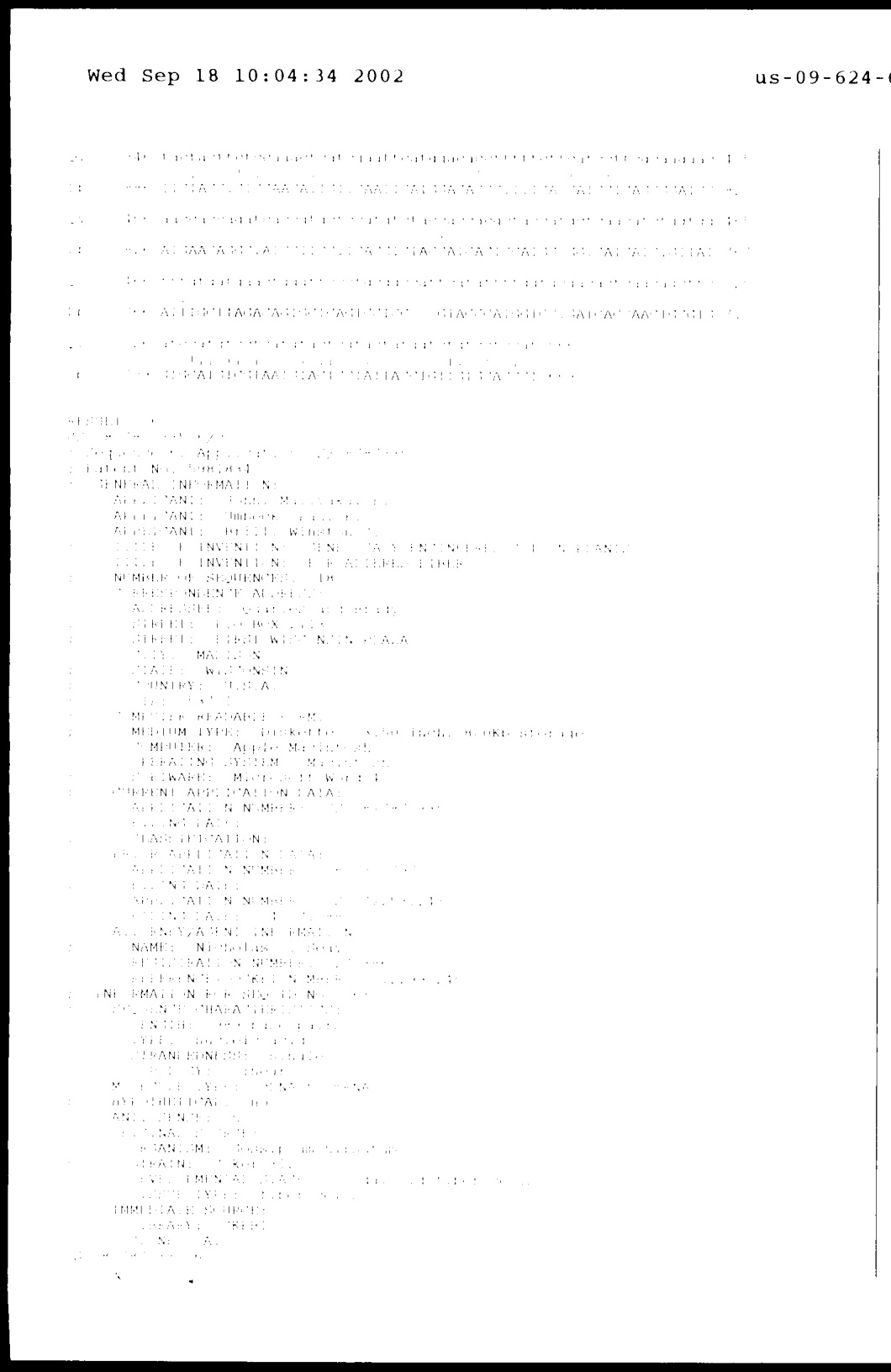
RESULT 2
US-07-885-970A-7/c
; Sequence 7, Application US/07885970A
; Patent No. 5495070
; GENERAL INFORMATION:

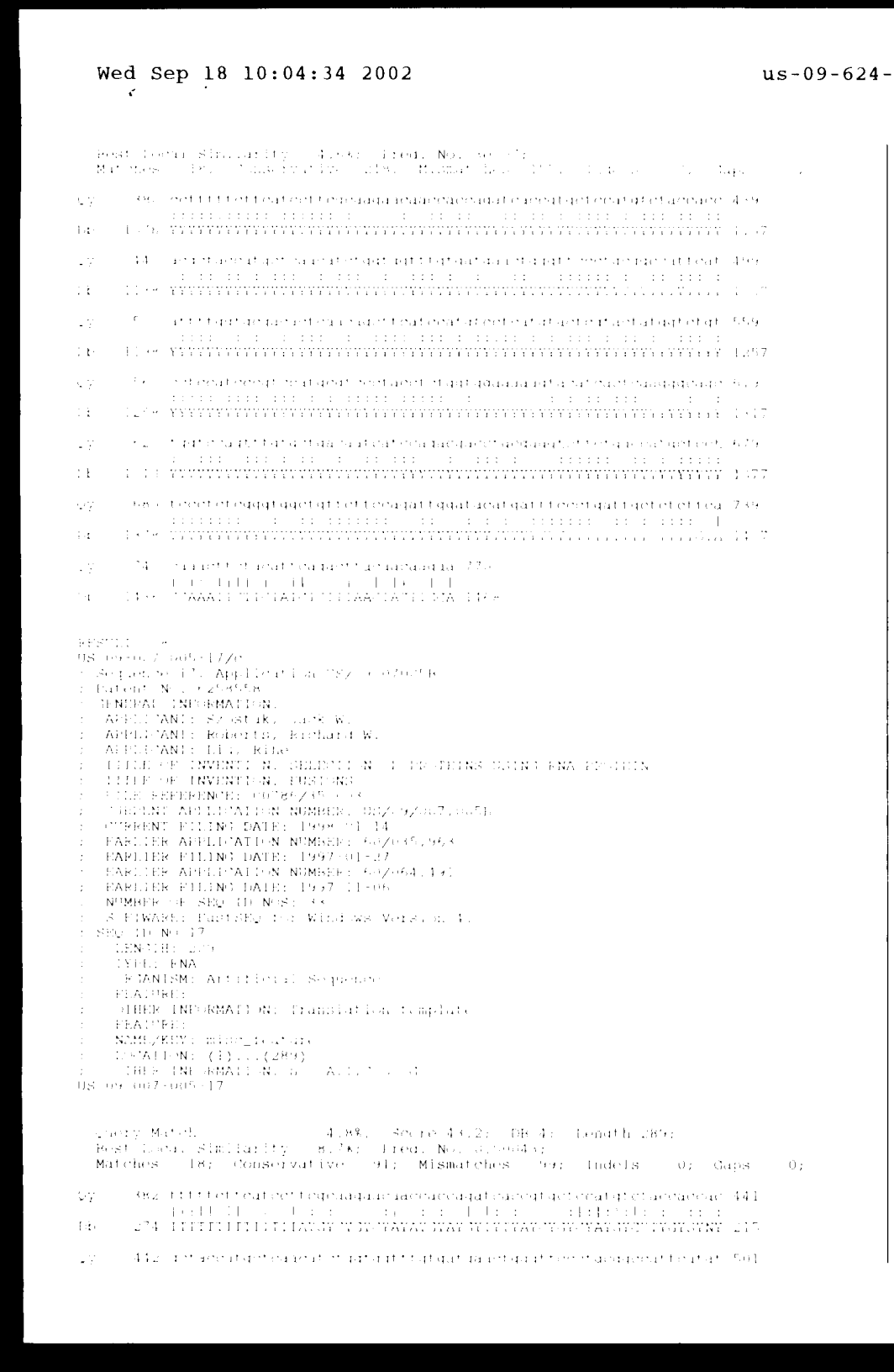


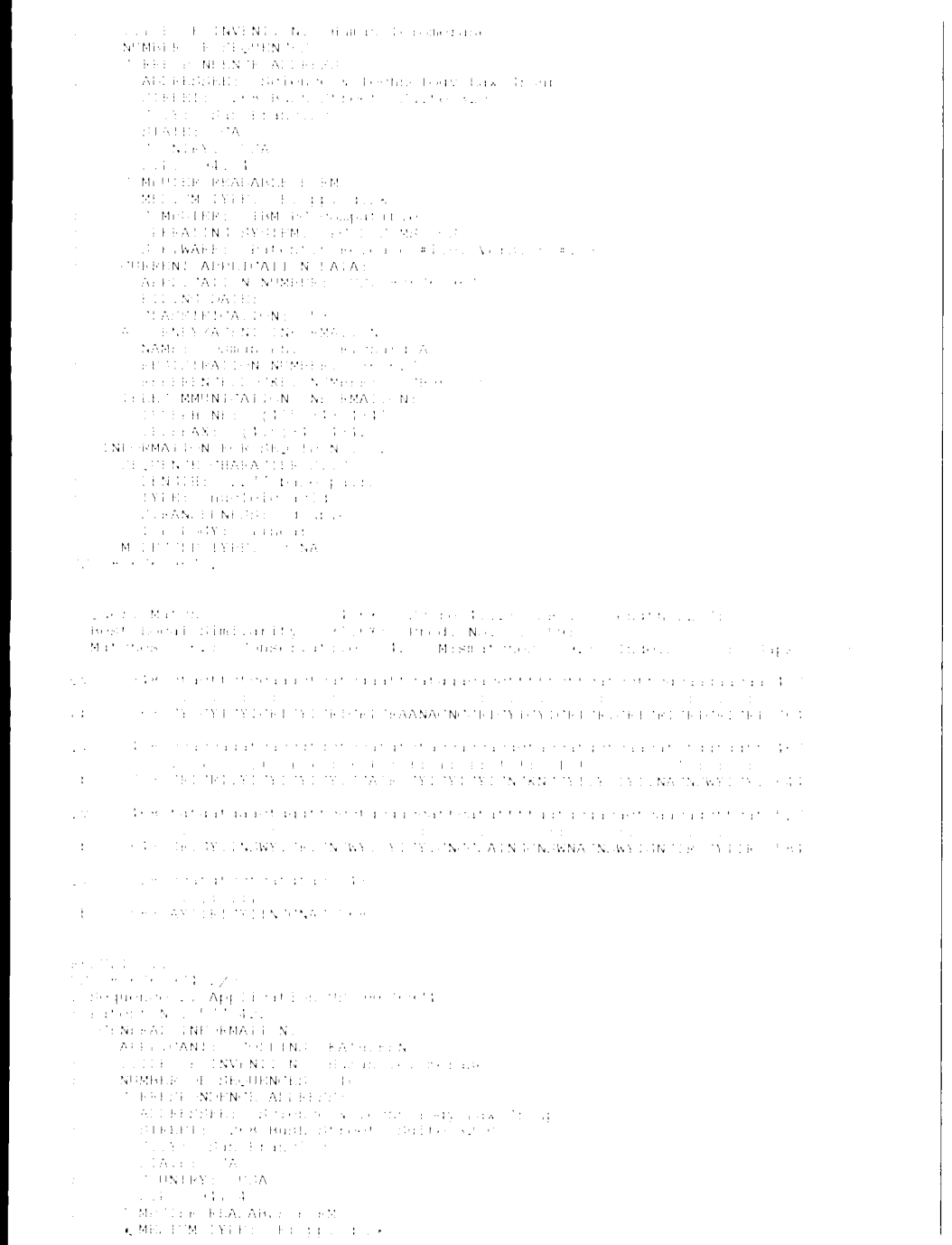


```
                 1 1111  1 1:111:1:1 11 11     111 : 1:1:
Db     709 GTGCATCCCCAACCTA*****************************  774


RESULT   4
US-08-540-797-6/c
; Sequence 6, Application US/08540797
; Patent No. 5597718
;  GENERAL INFORMATION:
;     APPLICANT:  John, Mariyakal E.
;     APPLICANT:  Umbeck, Paul F
;     APPLICANT:  Brill, Winston J.
;     TITLE OF INVENTION:  GENETICALY ENGINEERED COTTON PLANTS
;     TITLE OF INVENTION:  FOR ALTERED FIBER
;     NUMBER OF SEQUENCES:  38
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Quarles and Brady
;       STREET:  P.O BOX 2113
;       STREET:  FIRST WISCONSIN PLAZA
;       CITY:  MADISON
;       STATE:  WISCONSIN
;       COUNTRY:  U.S.A.
;       ZIP:  53701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette - 3.50 inch, 800Kb storage
;       COMPUTER:  Apple Macintosh
;       OPERATING SYSTEM:  Macintosh
;       SOFTWARE:  Microsoft Word 4.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/540,797
;       FILING DATE:  20-SEP-1995
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/617,239
;       FILING DATE:  21-NOV-90
;       APPLICATION NUMBER:  US 07/253,243
;       FILING DATE:  04-OCT-88
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Nicholas J. Seay
;       REGISTRATION NUMBER:  27,386
;       REFERENCE/DOCKET NUMBER:  1122990245
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  989 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA to mRNA
;     HYPOTHETICAL:  no
;     ANTI-SENSE:  no
;     ORIGINAL SOURCE:
;       ORGANISM:  Gossypium hirsutum
;       STRAIN:  Coker 312
;       DEVELOPMENTAL STAGE:  10 day old fiber cells
;       TISSUE TYPE:  fiber cells
;     IMMEDIATE SOURCE:
;       LIBRARY:  CKFB10
;       CLONE:  A11
US-08-540-797-6


  Query Match              7.0%;  Score 62.6;  DB 1;  Length 989;
  Best Local Similarity   58.4%;  Pred. No. 7.2e-10;
  Matches  129;  Conservative    0;  Mismatches   89;  Indels    3;  Gaps    1;

Qy      345 tactacttctccaaactatcaaattcatgcagcacctttcttcatcattcctcaaaaac 405
            | ||||  |||||||  |  | |||||||||||||||||  ||||  |||||||  ||     |
Db      886 TTCTACCTCTCCAAGATTGTTGAATTCATGGACAGTCTTTTGATCATGTTCAGCCGCATCC 827

Qy      406 gacaaccaaatacccatactcctatgtctaccaccaaactaccatgctcaacatctaqtqa 465
            |   |   | |  | | | | | ||||||| ||| | |   ||  ||| ||| || | |
Db      826 ATGAAGAGCCTATCCTTCCTTCAGCTCTACCATCACTCCATGGTGGTCATCATGTGTTAT 767
```

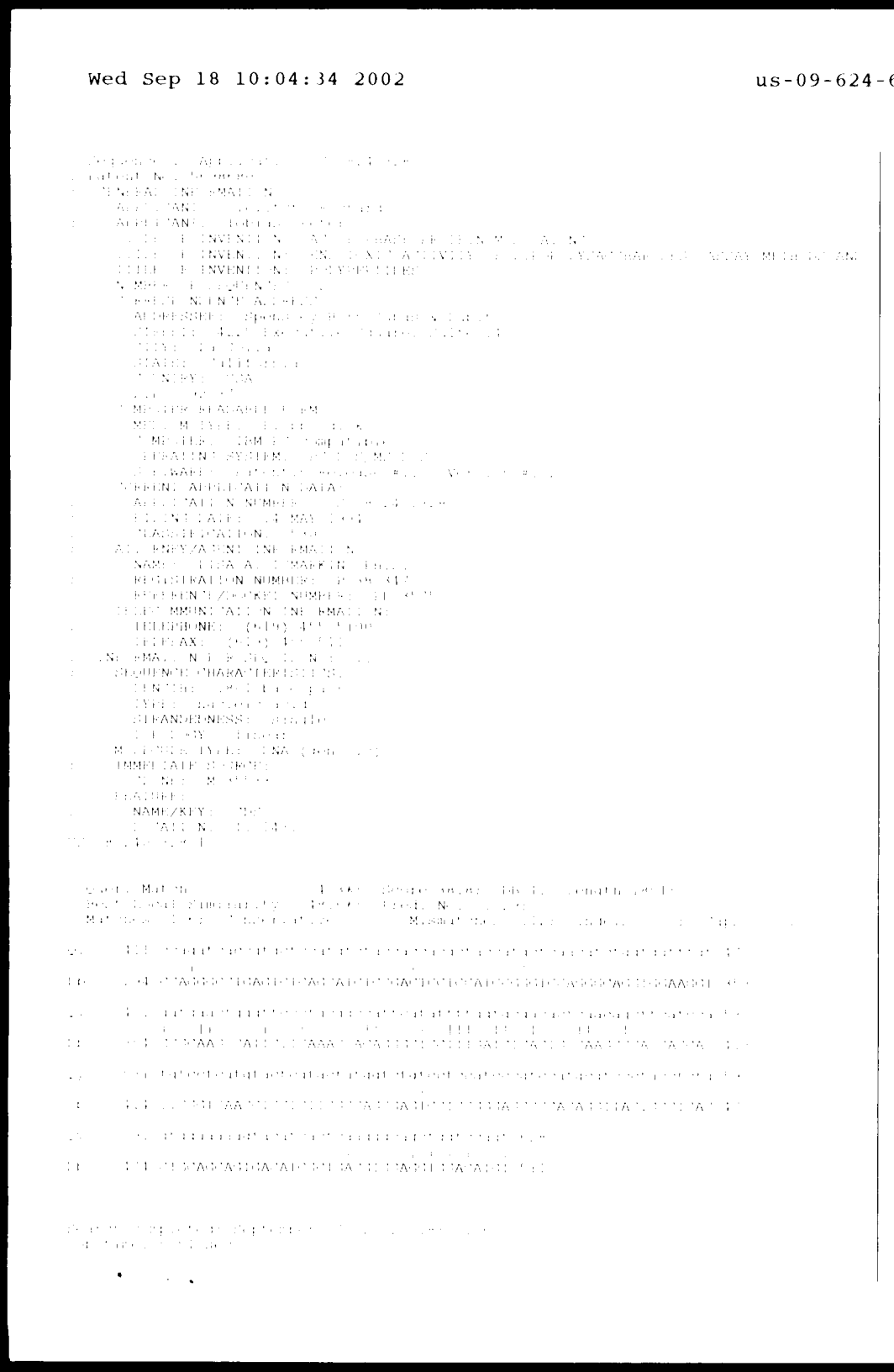
```
Qy      466 tttatgaatgaactgaqttcccctgcqaccattcatatttqatqcqacactcaacaqctc 525
            ||    | || | | | || | |  |    |    | | |    ||   |||| || |
Db      766 ATTTGCTTAGACAGTCCTCAGTCCTTC   GTACCATCCTCA TGATGAGCAACTGCGTG 710

Qy      526 atccatqtcctcatqtactcqtactatqqtctqtcctcat 566
            | |||| | ||||||| || ||     ||| | || |
Db      709 GTGCATGTCGTAATGTACTCCTATTACGTGTTGTGCATCCCT 669


RESULT   5
US-08-298-829-7/c
; Sequence 7, Application US/08298829
; Patent No. 5620882
;  GENERAL INFORMATION:
;     APPLICANT:  John, Mariyakal E.
;     TITLE OF INVENTION:  GENETICALLY ENGINEERING COTTON
;     TITLE OF INVENTION:  PLANTS FOR ALTERED FIBER
;     NUMBER OF SEQUENCES:  33
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Nicholas J. Seay, Quarles & Brady
;       STREET:  P.O. Box 2113, First Wisconsin Plaza
;       CITY:  Madison
;       STATE:  Wisconsin
;       COUNTRY:  USA
;       ZIP:  53701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  Microsoft Word
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/298,829
;       FILING DATE:  19-OCT-1994
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/885,970
;       FILING DATE:  18-MAY-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/617,239
;       FILING DATE:  21-NOV-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/253,243
;       FILING DATE:  04-OCT-1988
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Seay, Nicholas J.
;       REGISTRATION NUMBER:  27,386
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (608) 283-2478
;       TELEFAX:  (608) 251-5139
;   INFORMATION FOR SEQ ID NO:  7:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  989 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Gossypium hirsutum
;       STRAIN:  Coker 312
;       DEVELOPMENTAL STAGE:  10 day old fiber cells
;       TISSUE TYPE:  fiber cells
;     IMMEDIATE SOURCE:
;       LIBRARY:  CKFB10
;       CLONE:  A11
US-08-298-829-7


  Query Match              7.0%;  Score 62.6;  DB 1;  Length 989;
  Best Local Similarity   58.4%;  Pred. No. 7.2e-10;
  Matches  129;  Conservative    0;  Mismatches   89;  Indels    3;  Gaps    1;
```







```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (415)343-4341
;      TELEFAX: (415)343-4342
;  INFORMATION FOR SEQ ID NO: 2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 2277 base pairs
;      TYPE: nucleic acid
;      STRANDEDNESS: double
;      TOPOLOGY: linear
;    MOLECULE TYPE: cDNA
US-08-398-487-2

  Query Match             4.6%;  Score 41.2;  DB 2;  Length 2277;
  Best Local Similarity   41.6%;  Pred. No. 0.0659;
  Matches   42;  Conservative   42;  Mismatches   42;  Indels    0;  Gaps    0;

Qy        348 [illegible] 407
Db        [illegible]

Qy        408 [illegible] 467
Db        [illegible]

Qy        468 [illegible] 527
Db        [illegible]

Qy        528 [illegible] 54[illegible]
Db        [illegible]


RESULT 13
5245013
; Patent No. 5245013
;  APPLICANT: Ulevitch, Richard;Tobias, Peter
;  TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING ENDOTOXIC
;ACTIVITY OF LIPOPOLYSACCHARIDES, ASSAY METHODS AND POLYPEPTIDES
;  NUMBER OF SEQUENCES: 26
;  CURRENT APPLICATION DATA:
;    APPLICATION NUMBER: US/07/[illegible]
;    FILING DATE: 01-JUN-1989
;    APPLICATION NUMBER: [illegible]
;    FILING DATE: [illegible]
;    APPLICATION NUMBER: [illegible]
;    FILING DATE: [illegible]
;SEQ ID NO:5:
;    LENGTH: 1481
[illegible]

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length 1481;
  Best Local Similarity   48.6%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy        411 [illegible] 470
Db        [illegible]

Qy        471 [illegible]
Db        [illegible]

Qy        [illegible]
Db        [illegible]

Qy        [illegible]
Db        [illegible]


RESULT 14
US-07-847-562-1
; Sequence 1, Application US/07847562
; Patent No. 5310879
;  GENERAL INFORMATION:
;    APPLICANT: Ulevitch, Richard
;    APPLICANT: Tobias, Peter
;    TITLE OF INVENTION: ACUTE PHASE PROTEIN MODULATING ENDOTOXIC
;    TITLE OF INVENTION: ACTIVITY OF LIPOPOLYSACCHARIDES, ASSAY METHODS AND
;    TITLE OF INVENTION: POLYPEPTIDES
;    NUMBER OF SEQUENCES: 2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Spensley Horn Jubas & Lubitz
;      STREET: 4225 Executive Square, Suite 1400
;      CITY: La Jolla
;      STATE: California
;      COUNTRY: USA
;      ZIP: 92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/07/847,562
;      FILING DATE: 19920[illegible]
;      CLASSIFICATION: 530
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Wetherell, Jr. Ph.D., John R.
;      REGISTRATION NUMBER: 31,678
;      REFERENCE/DOCKET NUMBER: FD-1859
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (619) 455-5100
;      TELEFAX: (619) 455-5110
;  INFORMATION FOR SEQ ID NO: 1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 1801 base pairs
;      TYPE: NUCLEIC ACID
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;    MOLECULE TYPE: DNA (genomic)
;    IMMEDIATE SOURCE:
;      CLONE: M 35533
;    FEATURE:
;      NAME/KEY: CDS
;      LOCATION: 1..1431
US-07-847-562-1

  Query Match             4.3%;  Score 38.8;  DB 1;  Length 1801;
  Best Local Similarity   48.6%;  Pred. No. 0.029;
  Matches  106;  Conservative    0;  Mismatches  112;  Indels    0;  Gaps    0;

Qy        411 ccagatcaccatgctccatatctacaccacagctaccatgctcaacatctgatgatttat 470
Db        294 [illegible] 353

Qy        471 [illegible] 530
Db        354 [illegible] 413

Qy        531 [illegible] 590
Db        414 [illegible] 473

Qy        591 [illegible] 628
Db        474 CTGCAGCAGTGACACCGCTGACTGGAGGTGGACATGT 511


RESULT 15
US-08-249-328-[illegible]
```



SeqSearch version 4.1
Copyright (c) 1993 - 2002 Compugen Ltd.

Matrix: nucleic search, using sw model

Run on: September 17, 2002, 04:01:36 ; Search Time 34.1.72 Seconds
(without alignments)
[illegible] Million cell updates/sec

Title: US-09-624-670-6
Perfect score: 900
Sequence: 1 atttaaaaattttaatgaatttc.....gttttaaaaaaaaaaaaaaa 900

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1479627 seqs, 974847542 residues

Total number of hits satisfying chosen parameters: 2959254

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_estfun:*
3: em_estin:*
4: em_estmu:*
5: em_est_ov:*
6: em_estpl:*
7: em_est_re:*
8: em_est_un:*
9: em_htc:*
10: ab_estba:*
11: ab_htc:*
12: ab_gss1:*
13: em_gss_hum:*
14: em_gss_mus:*
15: em_gss_pro:*
16: em_gss_vrt:*

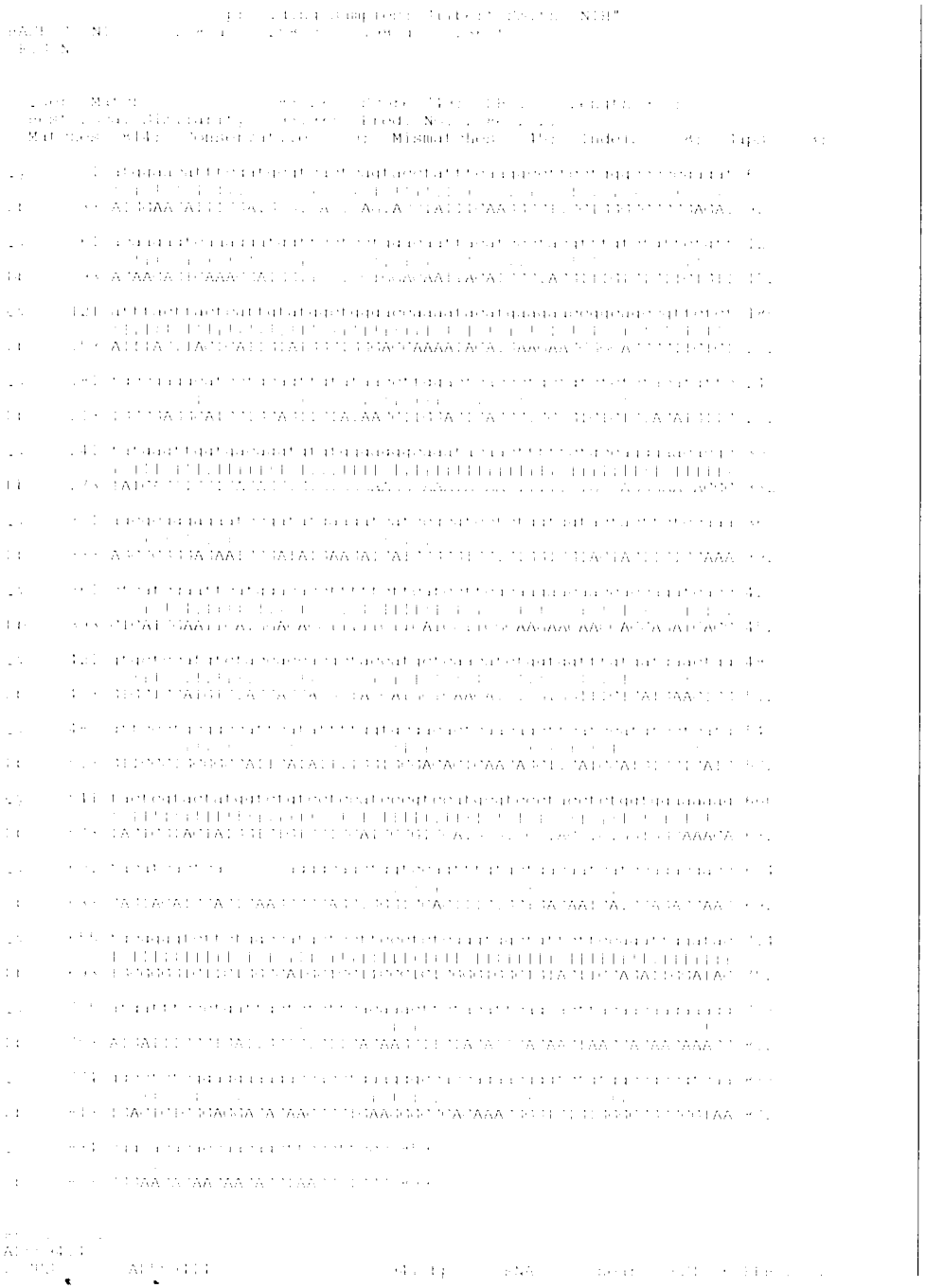
Prob. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

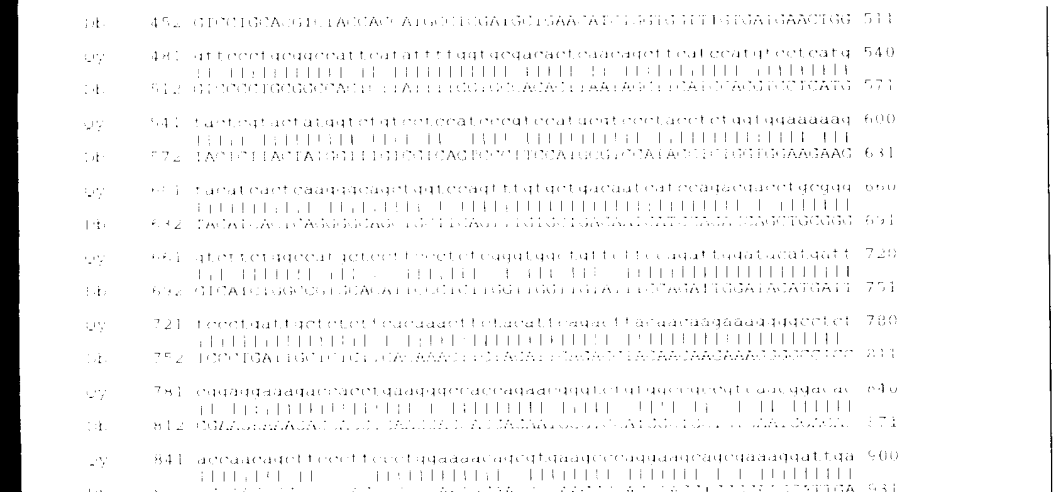
SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	900	100.0	900	10	BE161512	BE161512 601771006
2	900	100.0	900	9	BE080106	BE080106 BE1 B1062
3	900	100.0	900	9	AL548756	AL548756 AL548756
4	900	100.0	900	10	BE125375	BE125375 601087736
5	900	100.0	900	10	BE758566	BE758566 601402086
6	900	100.0	900	10	BE597218	BE597218 601250986
7	900	100.0	900	10	BE096205	BE096205 MR1-BT1069
8	900	100.0	900	9	BB661704	BB661704 BB661704
9	900	100.0	900	9	AL518949	AL518949 AL518949
10	900	100.0	900	9	BB621573	BB621573 BB621573
11	900	100.0	900	9	BB656883	BB656883 BB656883
12	900	100.0	900	9	BB643510	BB643510 BB643510
13	900	100.0	900	9	AW213134	AW213134 qg99d11.y
14	900	100.0	900	10	BE168059	BE168059 601776316
15	900	100.0	900	9	BB657364	BB657364 BB657364
16	900	100.0	900	10	BE251449	BE251449 602069505
17	900	100.0	915	10	BE144489	BE144489 601790129
18	900	100.0	978	9	AW740936	AW740936 wq16d18.y
19	900	100.0	900	10	BE668778	BE668778 602264725
20	900	100.0	900	10	BE837458	BE837458 601226161
21	900	100.0	951	9	AI749838	AI749838 qi58a08.y
22	900	100.0	410	10	BE944521	BE944521 UI-M-BB3-[illegible]
23	900	100.0	902	10	BE226416	BE226416 qz49b01.y
24	895	99.4	694	10	BE721563	BE721563 602695173
25	879	97.7	749	10	BE463309	BE463309 602204321

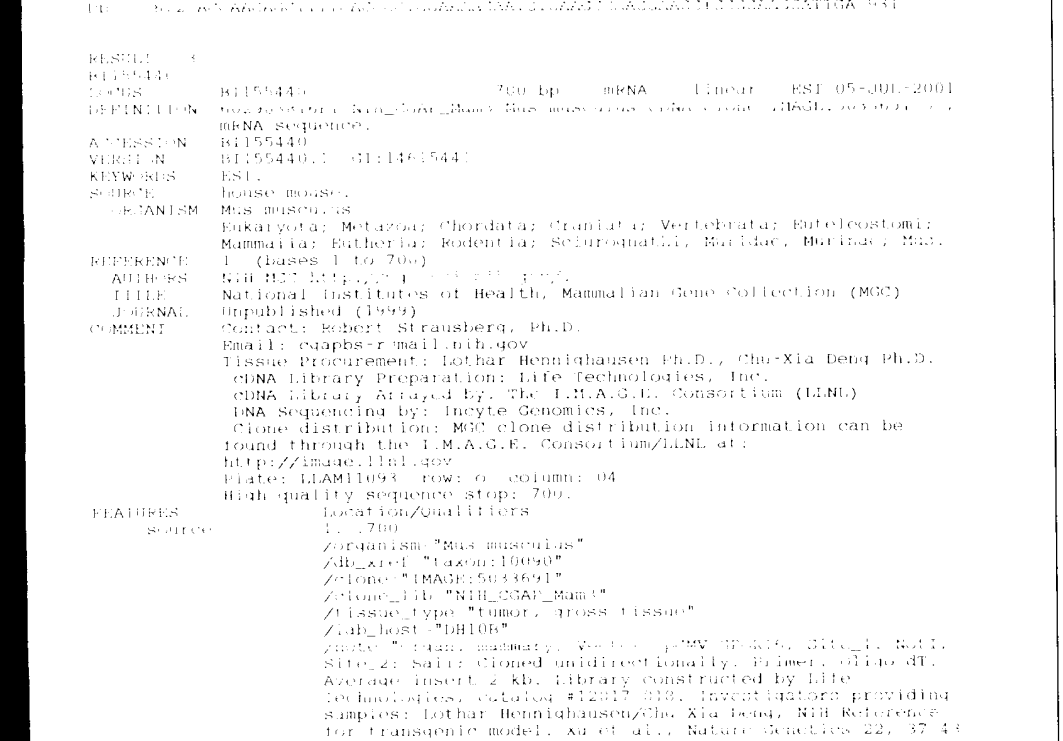
ALIGNMENTS

RESULT 1
BE161512
LOCUS BE161512 900 bp mRNA linear EST 29-SEP-2000
DEFINITION 601771006F1 NIH_MGC_Mus musculus cDNA clone IMAGE:[illegible], mRNA sequence.
ACCESSION BE161512
VERSION BE161512.1 GI:15955736
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 900)
AUTHORS [illegible]
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cmapbs@mail.nih.gov
Insert Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Fasta: library: 947 row: 1 column: 18
High quality sequence stop: 797.
FEATURES Location/Qualifiers
source 1..900
/organism="Mus musculus"
/strain="FVB/N"
/lib_xref="taxon:10090"
/clone="IMAGE:5872998"
/clone_lib="NIH_MGAI_Mam2"
/tissue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: vector dT. Library constructed by Life Technologies. Investigator



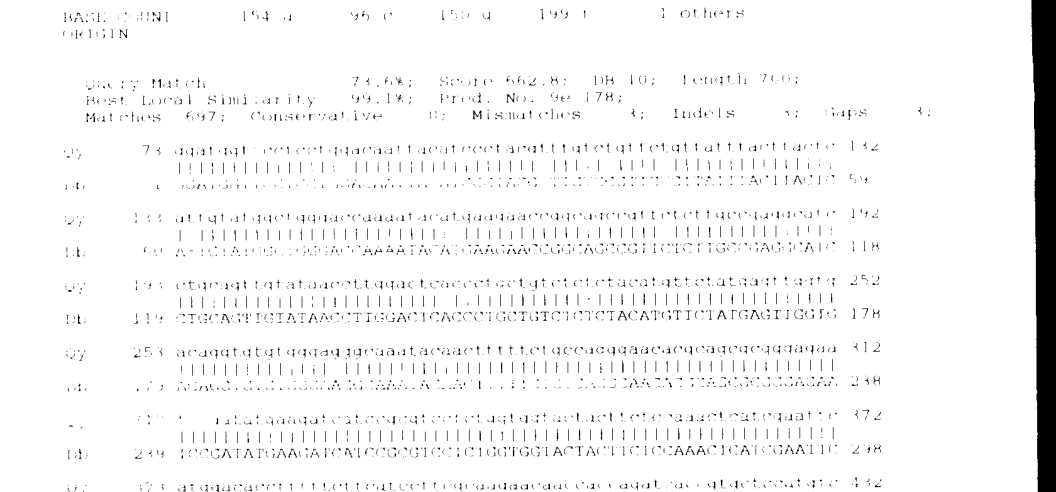


RESULT 3
BI155440
LOCUS BI155440 700 bp mRNA linear EST 05-JUL-2001
DEFINITION [illegible] mRNA sequence.
ACCESSION BI155440
VERSION BI155440.1 GI:14615441
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 700)
AUTHORS [illegible]
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11093 row: o column: 04
High quality sequence stop: 700.
FEATURES Location/Qualifiers
source 1..700
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5033691"
/clone_lib="NIH_CGAP_Mam1"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; Primer: Oligo dT;
Average insert 2 kb; Library constructed by Life
Technologies, Catalog #12017-018; Investigators providing
samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference
for transgenic model: Xu et al., Nature Genetics 22, 37-43
(1999); Note: this is a NCI CGAP Library."

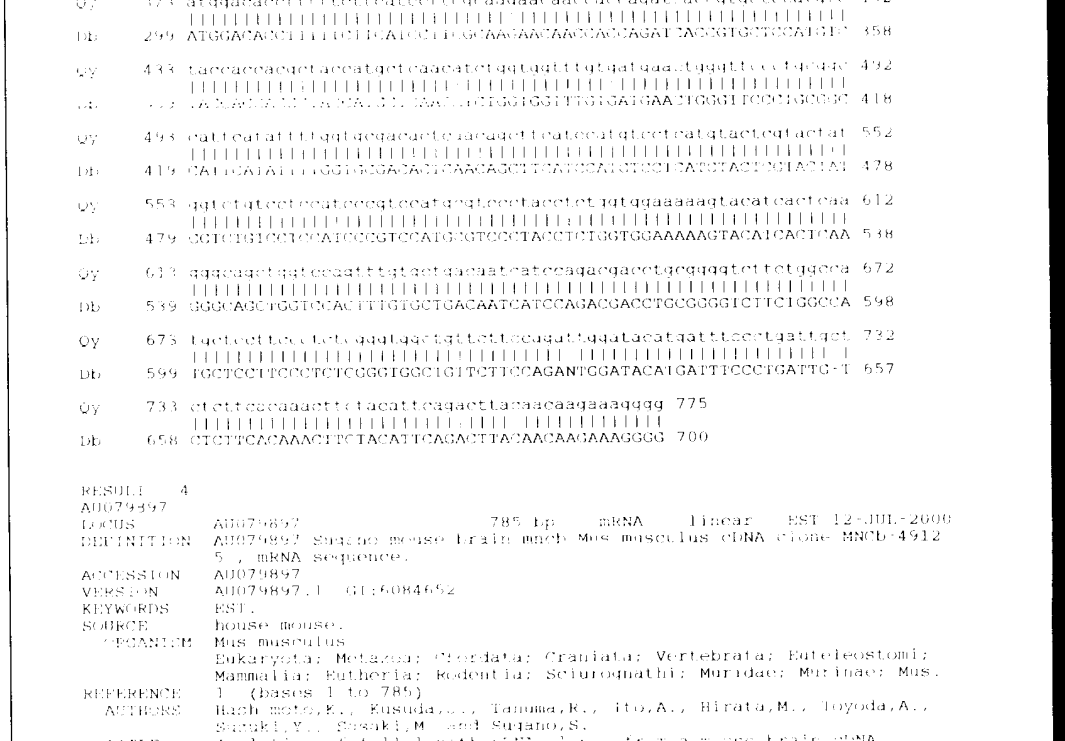


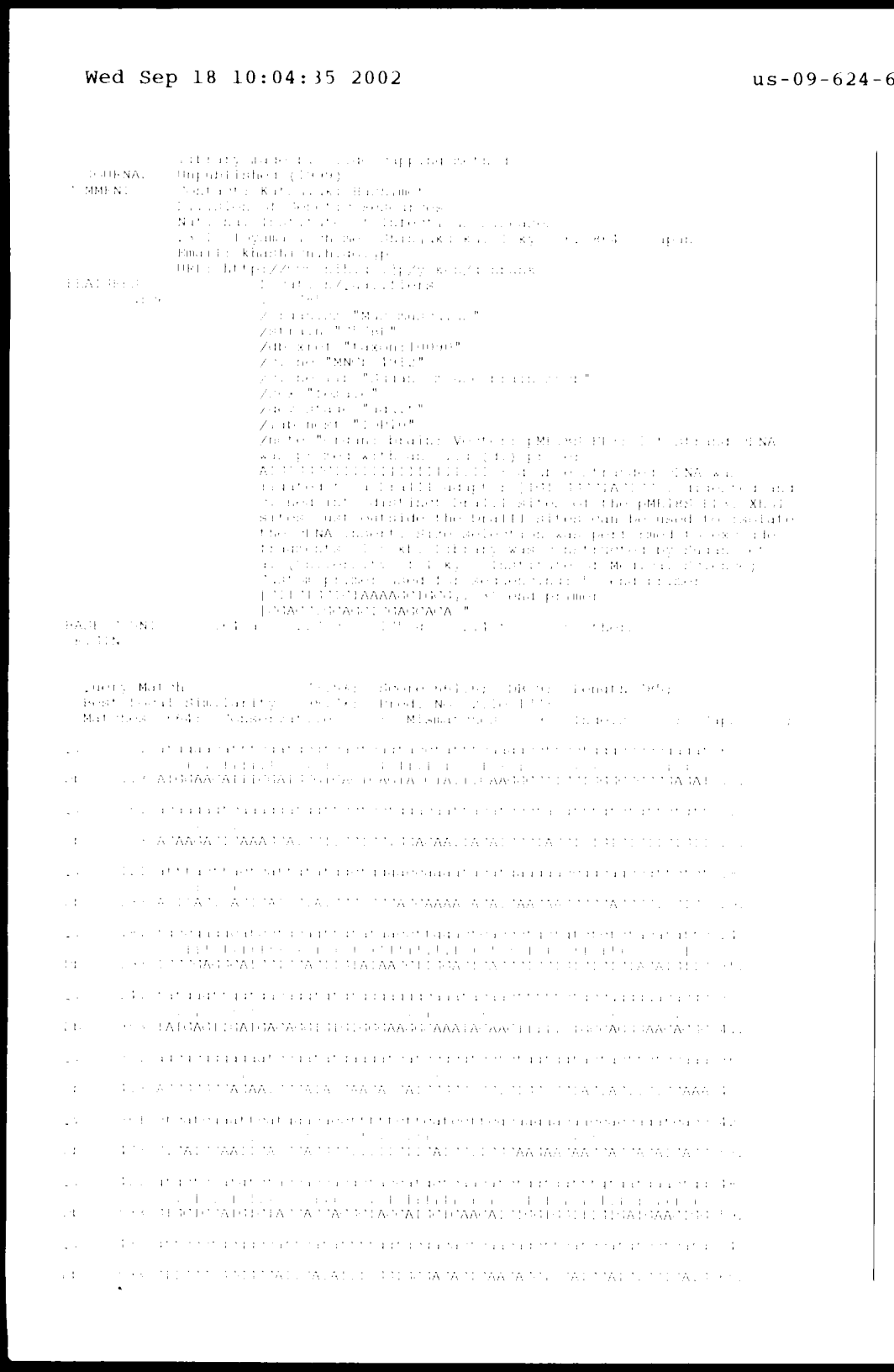
BASE COUNT 154 a 96 c 150 g 199 t 1 others
ORIGIN

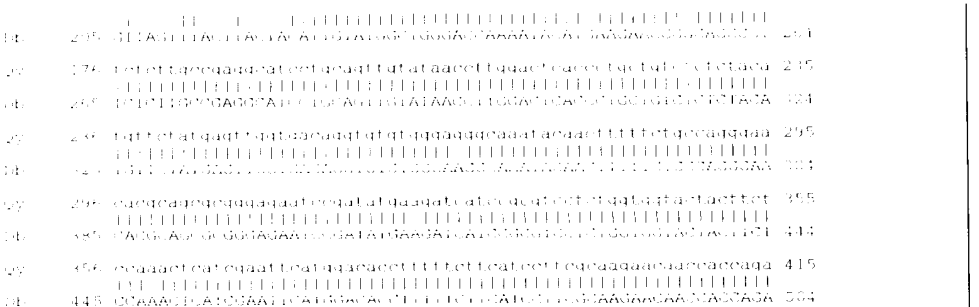
Query Match 73.6%; Score 662.8; DB 10; Length 700;
Best Local Similarity 99.1%; Prod. No. 9e-178;
Matches 697; Conservative 0; Mismatches 3; Indels 3; Gaps 3;



RESULT 4
AU079897
LOCUS AU079897 785 bp mRNA linear EST 12-JUL-2000
DEFINITION AU079897 Strain mouse brain mncb Mus musculus cDNA clone MNCb-4912
5 , mRNA sequence.
ACCESSION AU079897
VERSION AU079897.1 GI:6084652
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciuromorpha; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 785)
AUTHORS Nash mote,K., Kusuda,., Tanuma,K., Ito,A., Hirata,M., Toyoda,A.,
Sunaki,Y., Sasaki,M. and Sugano,S.
TITLE Isolation of full-length cDNA clones from a mouse brain cDNA







RESULT 6
BE545118
LOCUS BE545118 796 bp mRNA linear EST 09-AUG-2000
DEFINITION [illegible] mRNA sequence.
ACCESSION BE545118
VERSION BE545118.1 GI:9763763
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 796)
AUTHORS [illegible]
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM8770 row: I column: 10
High quality sequence stop: 644.
FEATURES location/qualifiers
source 1..796
/organism "Mus musculus"
/[illegible] "oyc/2]"
/db_xref "taxon:10090"
/clone "IMAGE:3595017"
/clone_lib "NCI_CGAP_Mam6"
/sex "female, virgin"
/tissue_type "infiltrating ductal carcinoma"
/dev_stage "5 months"
/lab_host "DH10B"
/note "Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigator
providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT 215 a 237 c 81 g 163 t
ORIGIN

Query Match 69.2%; Score 614.2; DB 10; Length 796;
Best Local Similarity 99.5%; Pred. No. 6.8e-164;
Matches 616; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 292 tttctgccaggaaacacgcaagcgcggcagaatccgatatgaagatcatccgcgctctg 351
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 TTTCTGCCAGGAACACGCAGCGCGGCAGAATCCGATATGAAGATCATCCGCGCTCTG 60

Qy 352 atggtactacttctccaaactcatcgaattcatggacacctttttcttcatcttcgccaa 411
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 GTGGTACTACTTCTCCAAACTCATCGAATTCATGGACACCTTTTTCTTCATCTTCGCCAA 120

Qy 412 caacaaccagcagatcacggtgcttcatgtctaccacgcacgctacgccatgctcaacatctg 471
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 CAACAACCAGCAGATCACGGTGCTTCATGTCTACCACGCACGCTACGCCATGCTCAACATCTG 180

Qy 472 atgatttatgatgaactgagttcctgcgccattcatatttgatgcgacactcaacaa 531
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 GTGCTTTGTGATGAACTGGGTTCCCTGCGGCCATTCATATTTAGGTGCGACACTCAACAG 240

Qy 532 cttcatccatgtcctcatgtactcgtactatgagtctgtcctccatcccgtccatgcatcc 591
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 241 CTTCATCCATGTCCTCATGTACTCGTACTATGGTCTGTCCTCCATCCCGTCCATGCGTCC 300

Qy 592 ctacctctagtgaagaaaatatcatgatcaaagcagctgagtccagtttttgtgctgacaat 651
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 301 CTACCTCTGGTGCAAAAAGTACATCACTCAAGGGCAGCTGGTCCAGTTAGTGCTGACAAT 360

Qy 652 catccacacgaccctccgggtcttctggccatgctccttccctctcgggtggctgttctt 711
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 361 CATCCAGACGACCTCCGGGGTCTTCTGGCCATGCTCCTTCCCTCTCGGGTGGCTGTTCTT 420

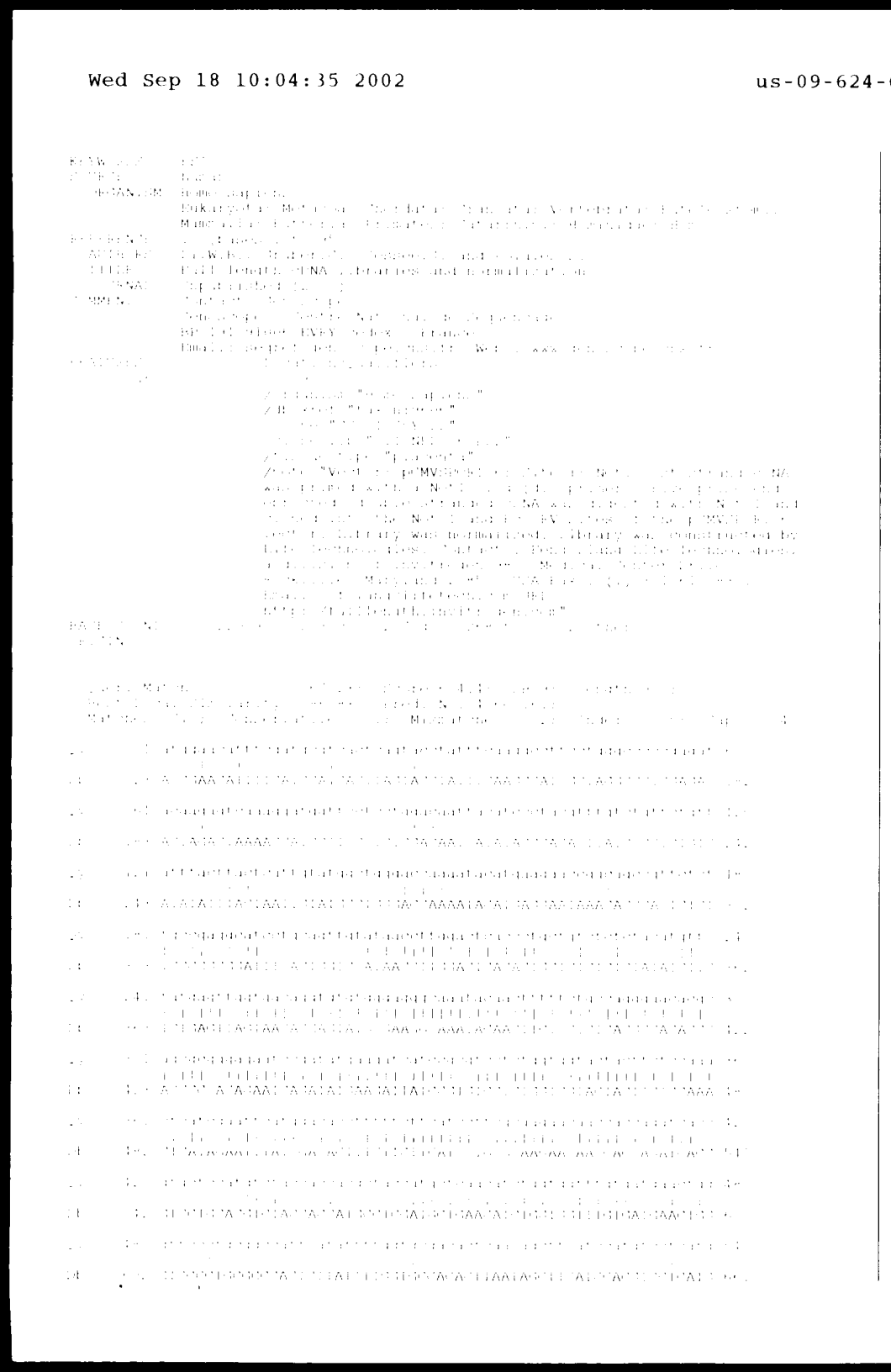
Qy 712 ccagattggatacatgatttccctgattgctctcttcacaaacttctacattcagactta 771
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 421 CCAGATTGGATACATGATTTCCCTGATTGCTCTCTTCACAAACTTCTACATTCAGACTTA 480

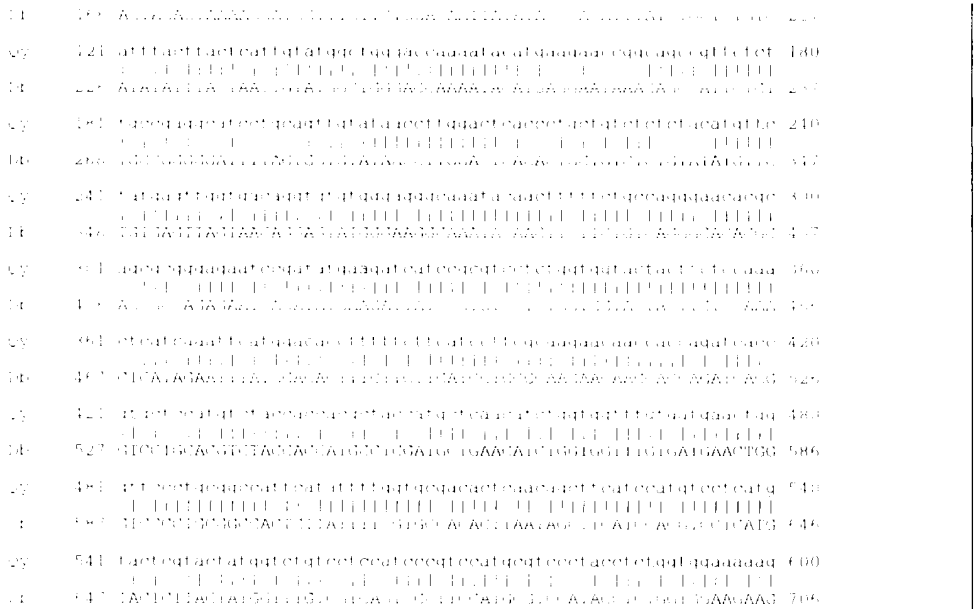
Qy 772 caaraaqaaaaqqqcctctcqqaqaaaaqaccacctqaaqqqccaccaqaaccqqtctat 831
||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 481 CAACAAGAAAAGGGGCTCTCGGAGCGAAAAGACCACCTGAAGGGCCACCAGAACCGGTCTGT 540

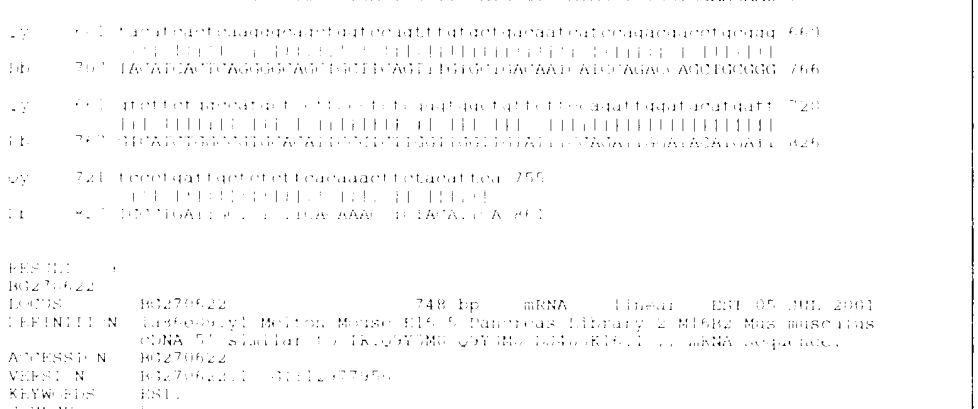
Qy 832 qqccqccqtcaacqqacacaccaacaqcttcccttcccttcqqaaacacaqccqtqaaqccca 891
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 541 GGCCGCCGTCAACGGACACACCAACAGCTTCCCTTCCCTGGAAACACAGCGTGAAGCCCAG 600

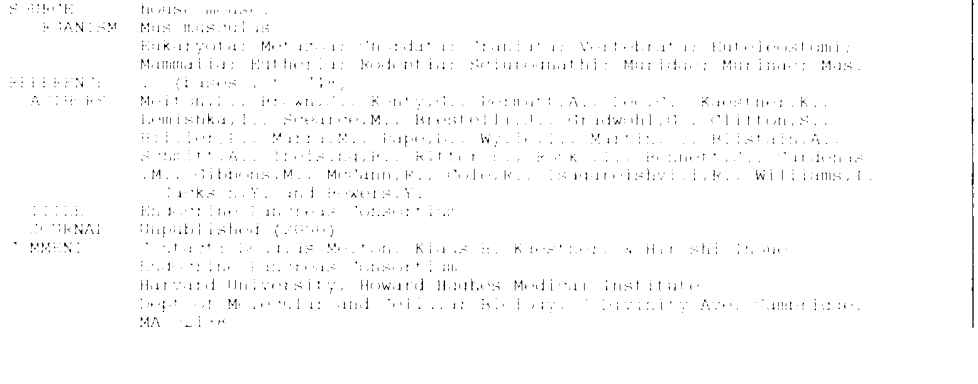
Qy 892 qaaqcaqcqaaaqqattqa 910
|||||| ||||||| ||||
Db 601 GAAGCAGCGAAAGGATTGA 619

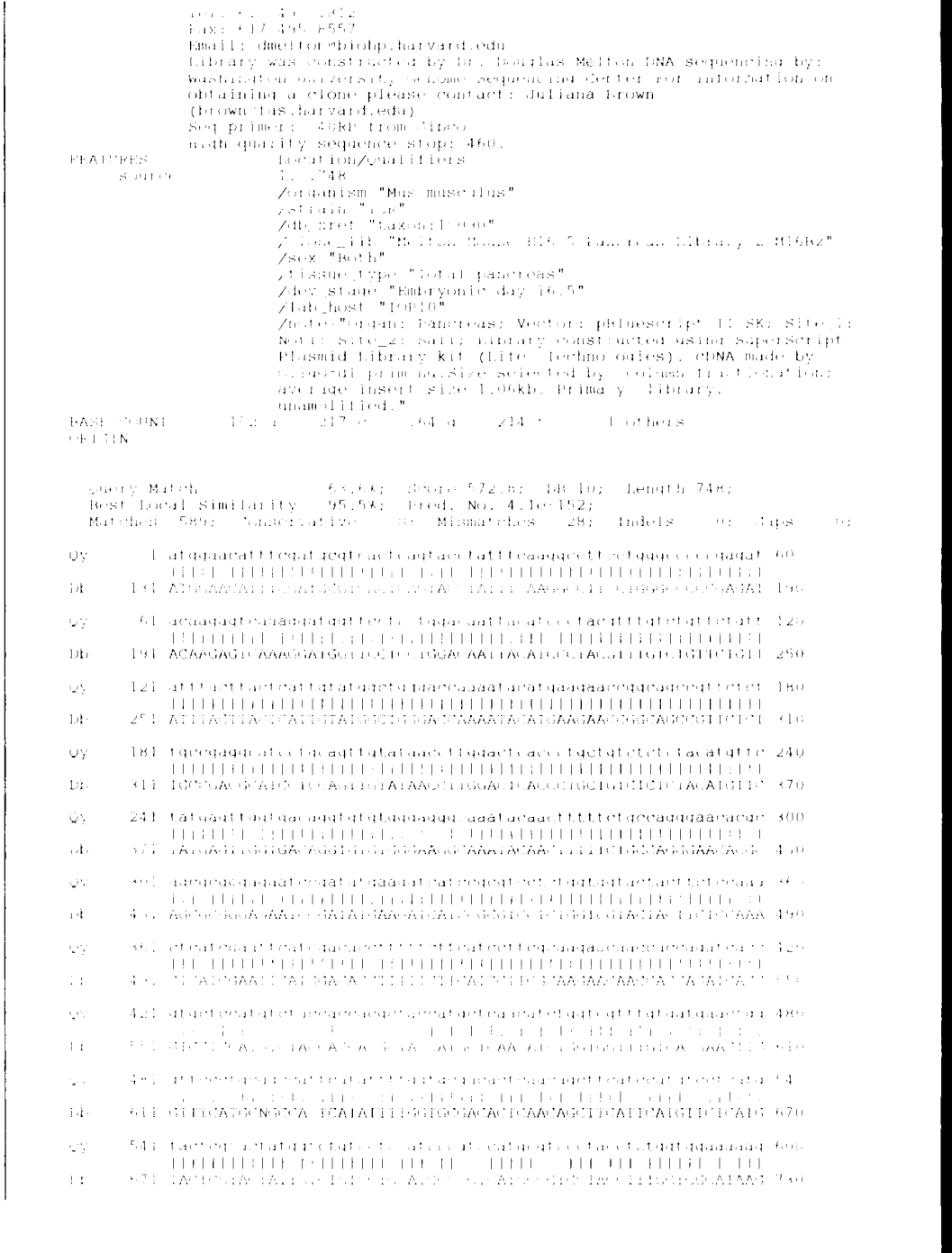
RESULT 7
AL553407
LOCUS AL553407 950 bp mRNA linear EST 16-FEB-2001
DEFINITION AL553407 LTI_NFL000_PE2 Homo sapiens cDNA clone CS0DI075YO22 5
prime, mRNA sequence.
ACCESSION AL553407
VERSION AL553407.1 GI:12893212

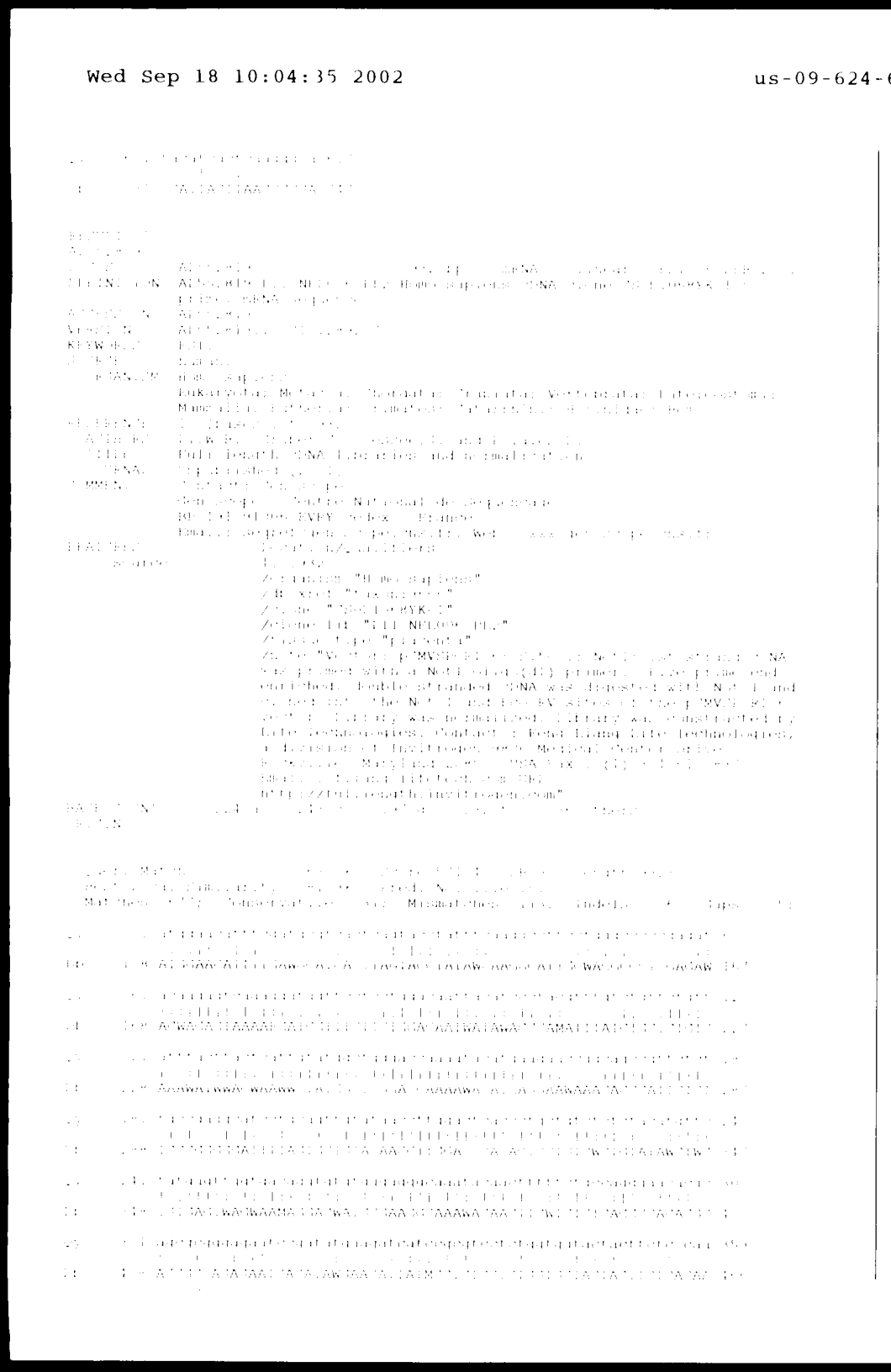


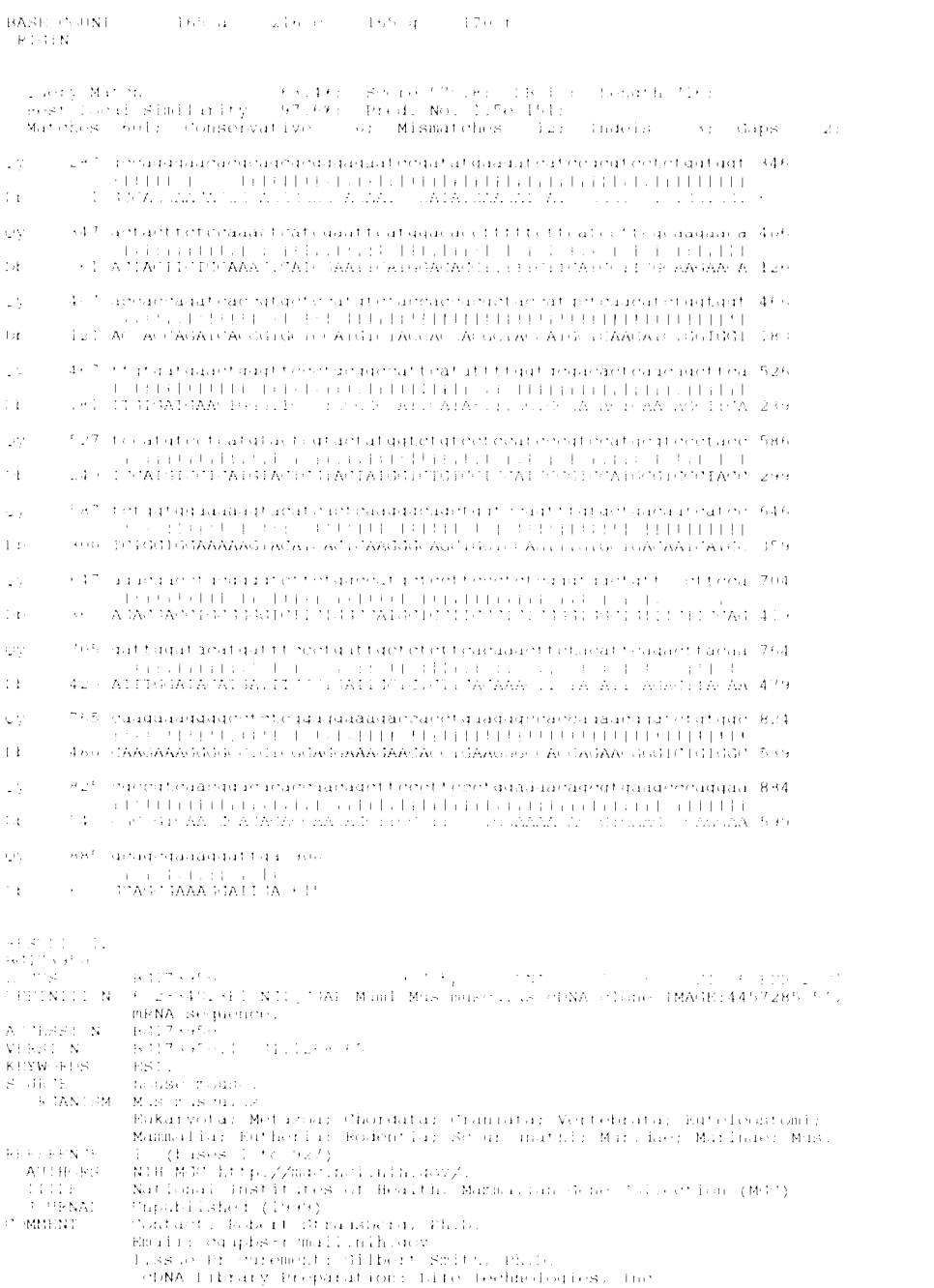




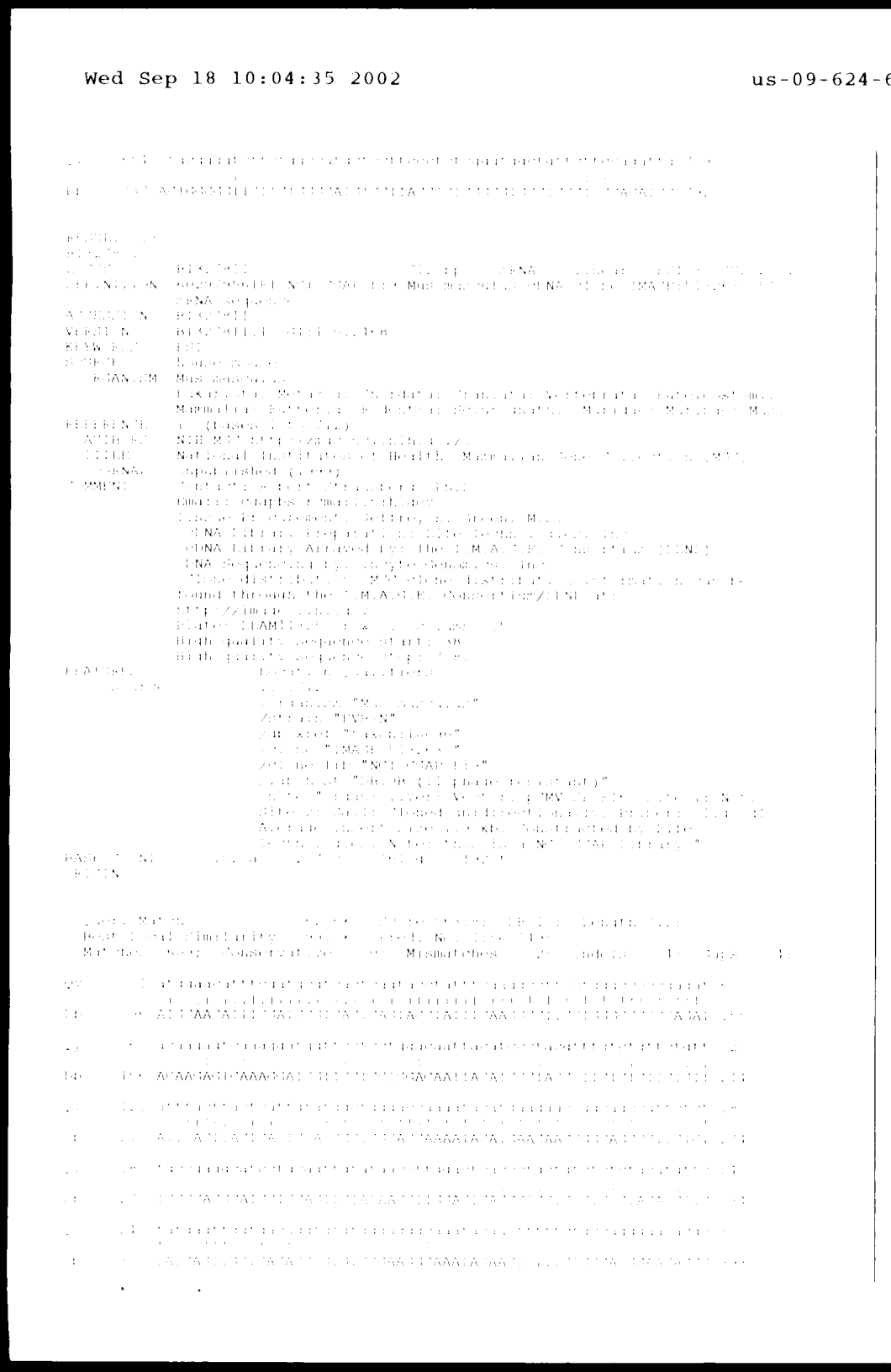


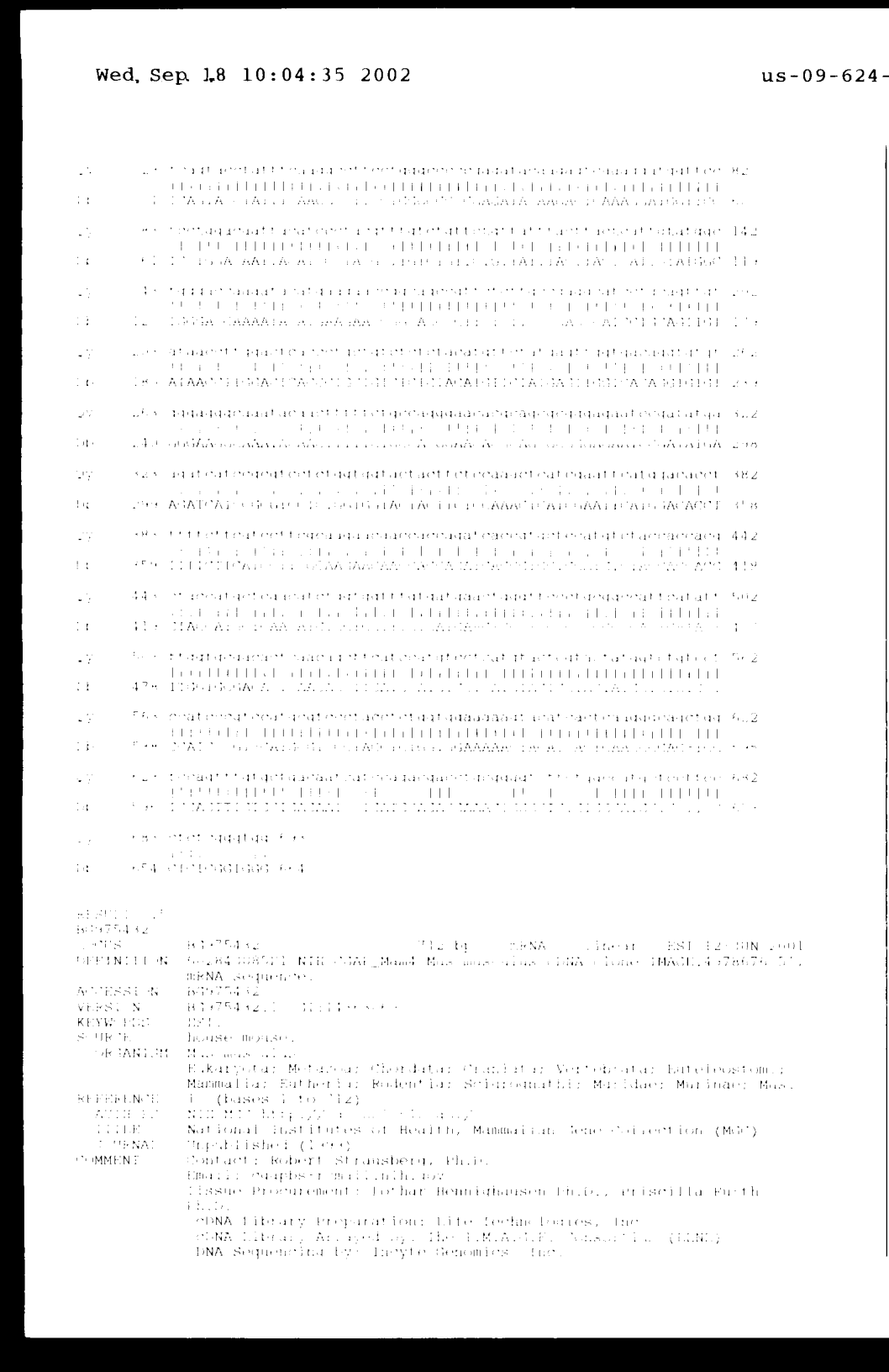






```
BASE COUNT       165 a    256 c    154 g    172 t
ORIGIN

  Query Match          64.4%;  Score 567.4;  DB 10;  Length 927;
  Best Local Similarity   92.6%;  Pred. No. 1.5e-150;
  Matches  663;  Conservative    0;  Mismatches   44;  Indels    5;  Gaps    6;

Qy      540 gttactcctactatgatcctatccctcgtcatgcatcctaactctgtt gtaaaa 598
            ||||||.|.| |||||||||i    | |.|.|.|.|.| | ||||||| |||| |
Db      617 GTACTCTCAGTTATCGTCGTCTCCATCTGTCTATGCGTCTGCAAGTCATCGGAAGAAAGA 676

Qy      599 agtacatcactcac  gatcaatgatgtcagttgt   gctacaatatcaggcagtc 653
            ||||||| | | ||  |||||| |||||||||||   |||||||    ||||||
Db      677 AGTACATCACATCAAAGGGCAGTTGATCTGGAGTCTGACATCATTCTCAGAAAGA 730
```





```
LOCUS       BI975432                 712 bp    mRNA    linear   EST 12-JUN-2001
DEFINITION  602864884R1 NIH_MGC_68 Mus musculus cDNA clone IMAGE:4376676 5',
            mRNA sequence.
ACCESSION   BI975432
VERSION     BI975432.1  GI:13715464
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 712)
  AUTHORS   NIH_MGC_http://mgc.nci.nih.gov/
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furth
            M.D.,
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: RAMI9925  Well: I  Vector: pCMV-SPORT6
            High quality sequence stop: 729.
FEATURES             Location/Qualifiers
     source          1..712
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4376676"
                     /clone_lib="NIH_MGC_68 Mus"
                     /tissue_type="Adult Liver Tissue"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6; Site_1: Not I; Site_2: Sall; Cloned unidirectionally, Primers: oligo dT. Average insert size: 2.5 kb. Library constructed by Life Technologies, catalog # 12018-016. Investigators providing samples: Lothar Hennighausen/Priscilla Furth, NIH. Reference for transgenic model: Li et al., Cell Growth and Differentiation 1, 541 (1990). NOTE: This is a NCI-CGAP Library."
BASE COUNT       154 a    209 c    192 g    156 t
ORIGIN
```

```
  Query Match             92.5%;  Score 960.4;  DB 10;  Length 712;
  Best Local Similarity   91.7%;  Pred. No. 1.4e-248;
  Matches  654;  Conservative    0;  Mismatches   16;  Indels    6;  Gaps    3;
```

GenCore version 4.5
Copyright (c) 1993 - 2002 Compugen Ltd.

[illegible] September [illegible] ; Search Time [illegible] Seconds
(without alignments)
[illegible] Million cell updates/sec

[illegible]
[illegible]
Sequence: [illegible]

Scoring table: IDENTITY NUC
Gapop [illegible]

Searched: 1370427 seqs, 670814752 residues

Total number of hits satisfying chosen parameters: 2747414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match [illegible]%
Maximum Match [illegible]%
Listing first [illegible] summaries

Database: EST:*
1: em_estfun:*
2: em_estfun2:*
3: em_estfun3:*
4: em_estmou:*
5: em_estrod:*
6: em_estpln:*
7: em_estrov:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_mus:*
15: em_gss_pln:*
16: em_gss_rod:*

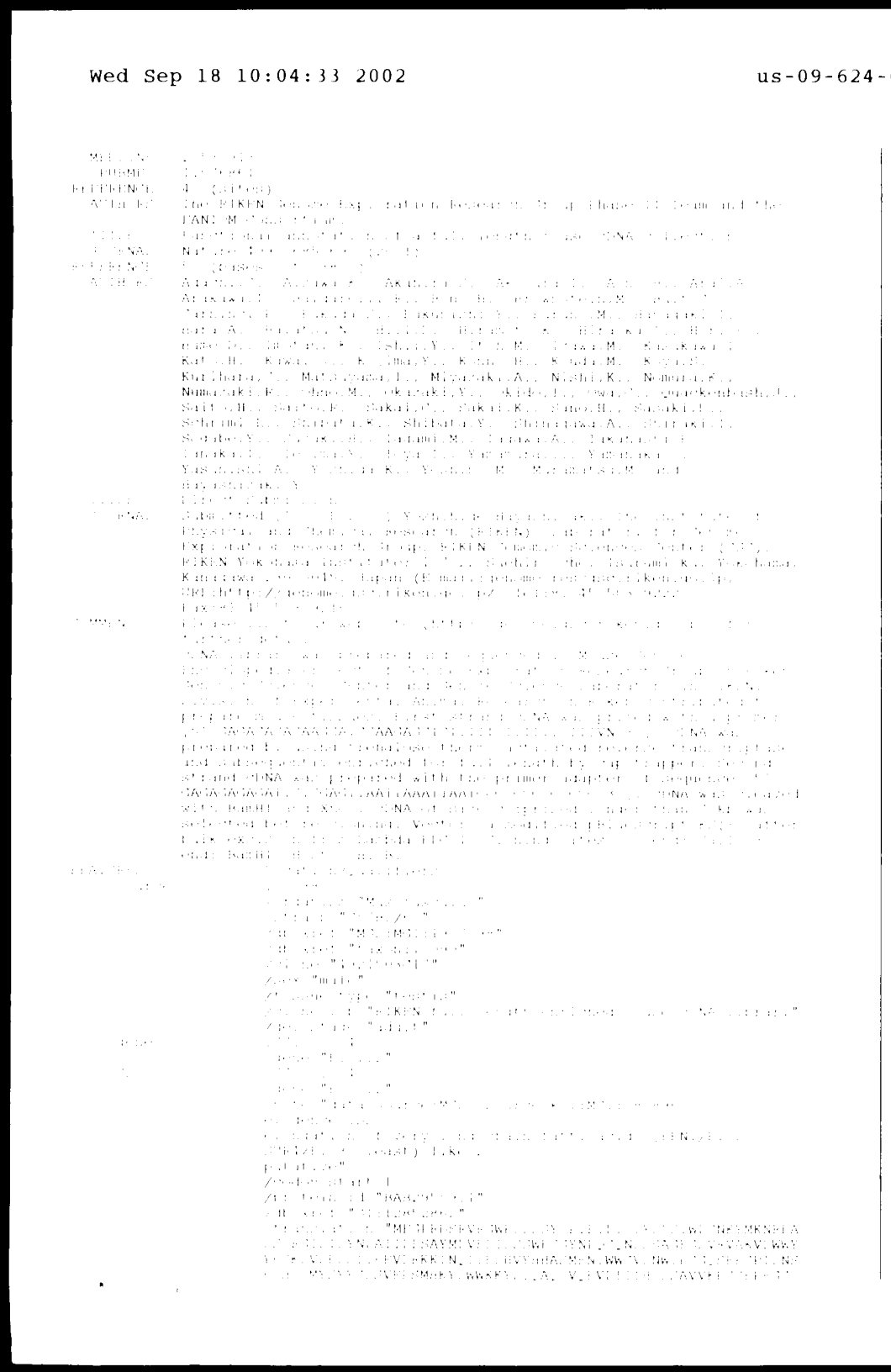
Total No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

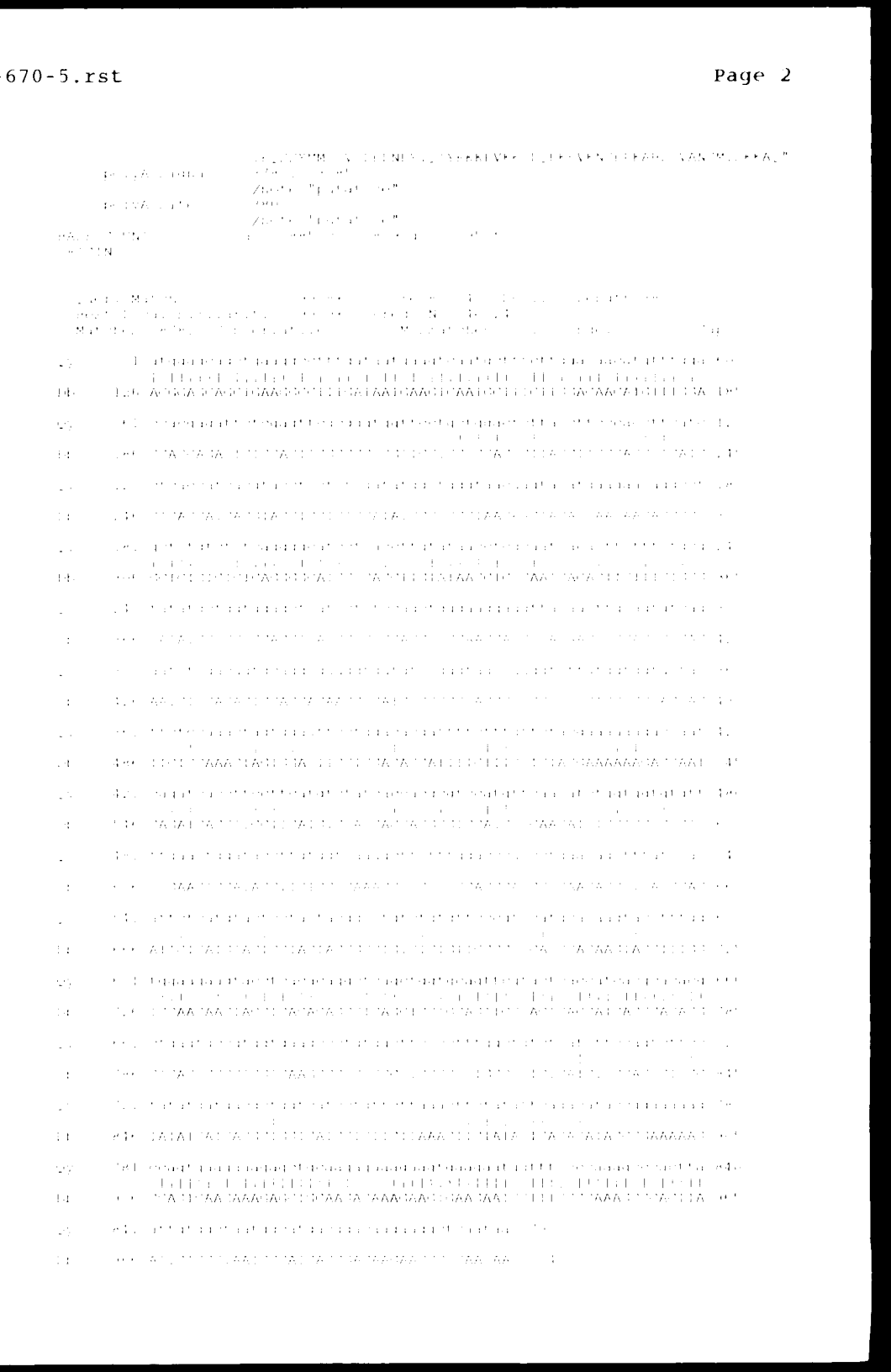
SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	8774.4	96.8	9061	12	AK014803	AK014803 Mus muscu
2	7481.2	93.0	1179	10	BM61475	BM61475 AZ2N2B3C
3	6056	88.9	661	9	BB654040	BB654040 BB654040
4	5796.8	87.8	643	9	BB616155	BB616155 BB616155
5	5662	83.9	645	9	BB641301	BB641301 BB641301
6	5544	81.9	623	9	BB642836	BB642836 BB642836
7	5528	80.1	589	9	AI172603	AI172603 UI-R-C2p-
8	5242.4	87.2	527	10	BE949059	BE949059 UI-M-BH3
9	4960.2	73.9	524	9	AW38918	AW38918 UI-M-BH3
10	4900	85.7	504	10	BE461042	BE461042 UI-M-C1p
11	4900	86.3	441	9	AI595458	AI595458 mi43b09.y
12	4864	84.8	464	9	AI142812	AI142812 mi19c07.x
13	4752.4	87.8	441	10	BE479061	BE479061 UI-M-BH5
14	4700	84.6	468	9	BE278465	BE278465 601463149
15	4696.8	84.1	468	10	BI426858	BI426858 pindmpk0
16	4696.8	84.1	468	10	BI426859	BI426859 pindmpk0
17	4696.8	84.1	468	10	BI426860	BI426860 pindmpk0
18	4696.8	84.1	468	10	BI426861	BI426861 pindmpk0
19	4684.2	88.0	435	9	AI853990	AI853990 mi11c04.y
20	4684	87.9	435	9	AA061889	AA061889 mi14b09.s
21	4590.4	86.3	441	9	AI550414	AI550414 AI550414
22	4512.4	85.5	861	9	AI514872	AI514872 AI514872
23	4511	85.3	958	9	AI554407	AI554407 AI554407
24	4508.2	85.2	905	10	BI691367	BI691367 601444742
25	4501.2	84.6	708	10	BI559440	BI559440 602003814
26	4496.6	84.2	525	10	BG814561	BG814561 ba14067.y
27	4494.4	84.1	785	9	AI173897	AI173897 AI173897
28	4490.8	84.0	661	9	BG80116	BG80116 BG80116
29	4481.4	84.0	802	9	AI516234	AI516234 AI516234
30	4477.2	83.8	760	9	AI548796	AI548796 AI548796
31	4476.8	83.8	952	9	AI552813	AI552813 AI552813
32	4468	83.7	687	10	BE577628	BE577628 [illegible]
33	4466	83.6	787	10	BI597218	BI597218 [illegible]
34	4466	83.6	884	10	BI760566	BI760566 [illegible]
35	4463.8	83.6	748	10	BG270022	BG270022 ba85009.y
36	4454.2	83.4	875	10	BI851179	BI851179 602379927
37	4454.2	83.4	810	10	BG251449	BG251449 602369905
38	4447.2	83.3	625	10	BM148219	BM148219 tp12r0039
39	4447	83.3	796	10	BE535118	BE535118 601241235
40	4446.6	83.3	682	10	BF096205	BF096205 MR3-BT009
41	4446.6	83.3	712	10	BI427811	BI427811 602979561
42	4445.4	83.3	653	10	BI559532	BI559532 602852873
43	4445.4	83.3	807	10	BE180425	BE180425 601756776
44	4444.6	83.3	755	10	BG120776	BG120776 602454420
45	4444.2	83.3	712	10	BG975432	BG975432 602843065

ALIGNMENTS

RESULT 1
AK014803
LOCUS AK014803 863 bp mRNA linear BSS 17-JAN-2002
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4921508C17, contains of very long chain fatty acids (FEN1/ [illegible]) [illegible] 5, full insert sequence.
ACCESSION AK014803
VERSION AK014803.1 GI:12952861
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain: [illegible]) adult male testis cDNA to mRNA.
clone lib:RIKEN full-length enriched mouse cDNA library
clone:4921508C17.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
REFERENCE 1 (sites)
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99277258
PUBMED [illegible]
REFERENCE 2 (sites)
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of Cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 (sites)
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,K., Matsubara,K., Hohsaka,R., Ikegami,T., Fukuda,K., Fujiwake,S., Inoue,K., Isogai,Y., Kasuga,M., Shiraki,T., Watahiki,M., Yoda,Y., Ishikawa,T., Lee,H., Tanaka,T., Matsuura,S., Baldarelli,R., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384 format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 8 (11), 1277-1291 (2000)





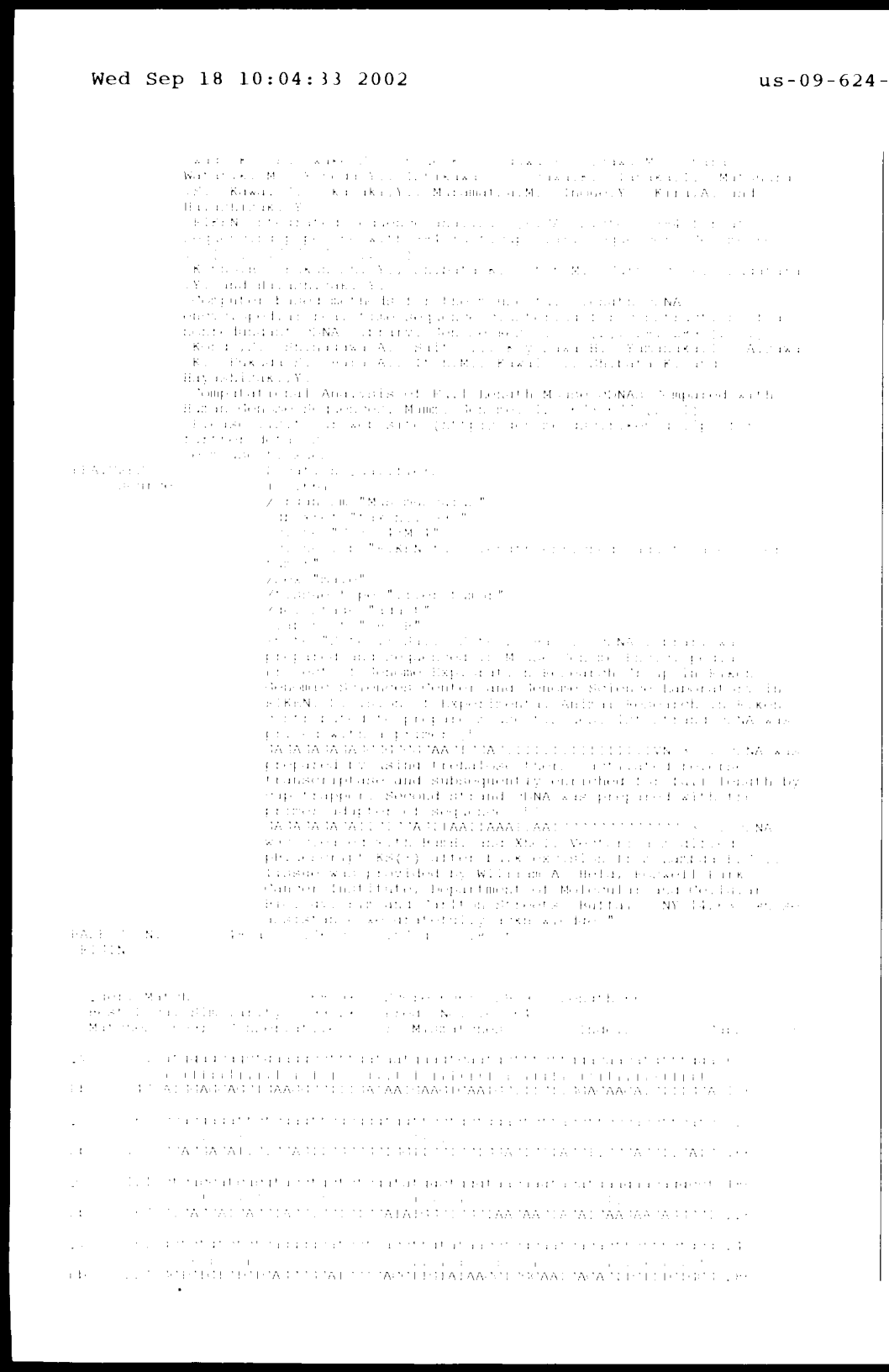
```
RESULT    2
BM461375
LOCUS       BM461375                1150 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6414097 NCI_CGAP_Ov44 Mus musculus cDNA clone
            IMAGE:5504127 5', mRNA sequence.
ACCESSION   BM461375
VERSION     BM461375.1  GI:18510415
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1150)
  AUTHORS   NIH MGC Project
  TITLE     National Institutes of Health Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Aaron Hsueh
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IMAM1214  row: p  column: 16
            High quality sequence start: 11
            High quality sequence stop: 664.
FEATURES             Location/Qualifiers
     source          1..1150
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5504127"
                     /clone_lib="NCI_CGAP_Ov44"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="organ: ovary, PMSG-treated; Vector:
                     pCMV.SPORT6.ccdb; Site_1: EcoRV; Site_2: NotI; Cloned
                     unidirectionally. Primer: oligo dT. Average insert size
                     2.2 kb. Library constructed by Life Technologies. Note:
                     this is a NCI_CGAP library."
BASE COUNT      288 a     327 c     244 g     291 t
ORIGIN
```

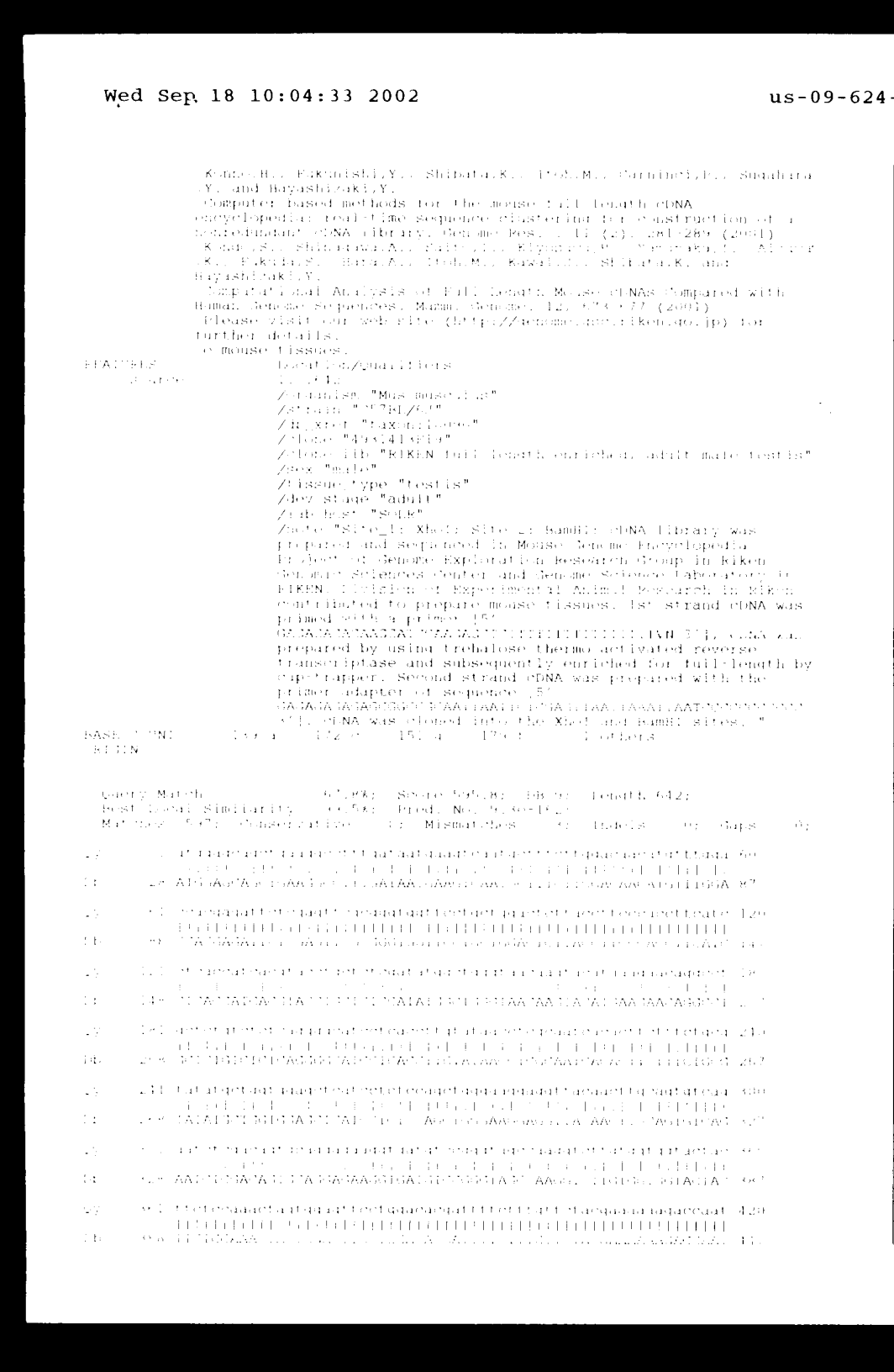
Query Match 90.0%; Score 791.2; DB 10; Length 1150;
Best Local Similarity 97.9%; Prod. No. 2.4e-216;
Matches 854; Conservative 9; Mismatches 14; Indels 5; Gaps 3;

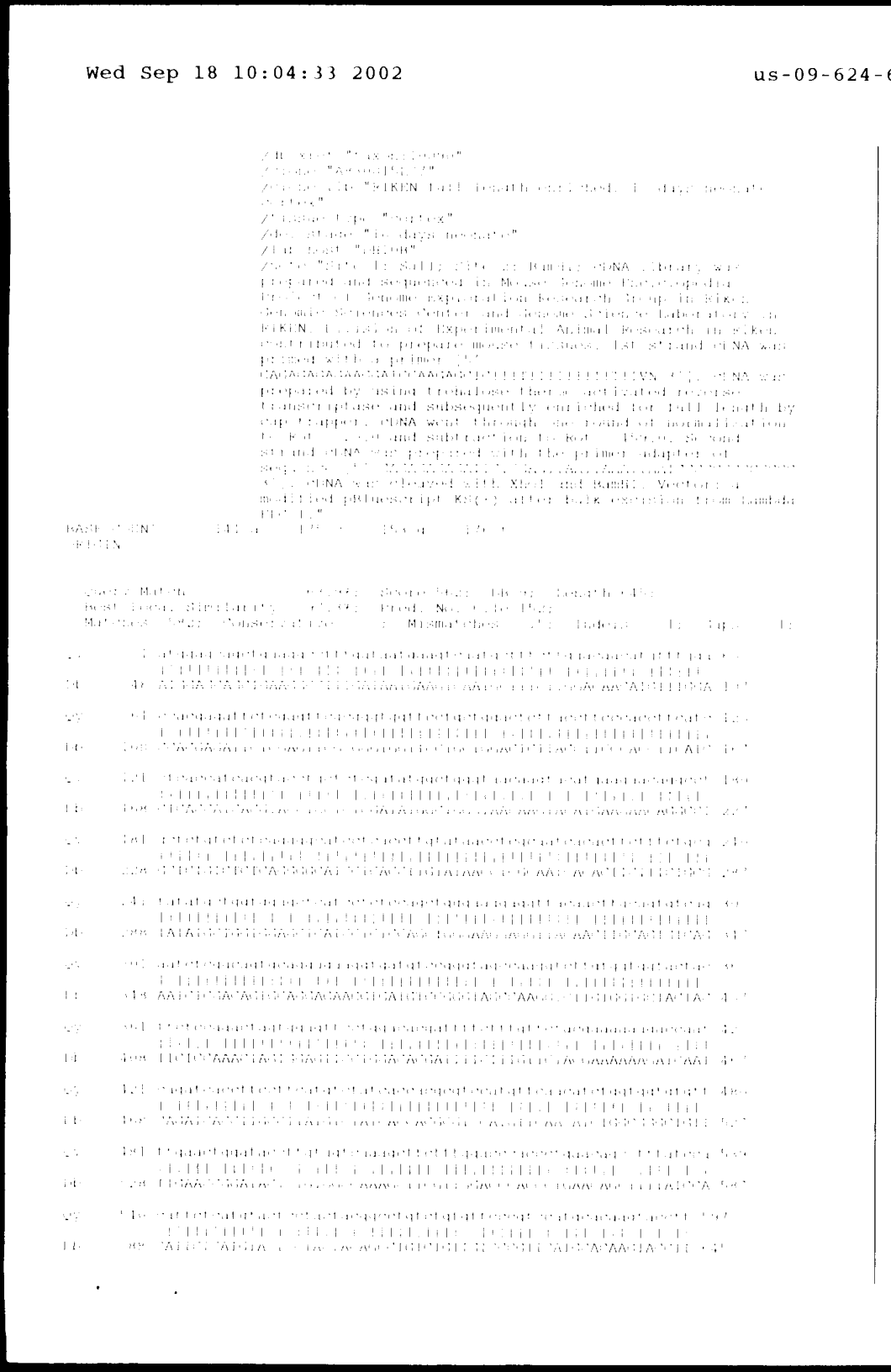
```
Qy      55 caatgcttttcttgaacaaatatt-aaaacaaaaatttcaaattcaaaaattttct 112
           ||||||||||||||||||||||||| |||||||||||||||||||||||||||||||
Db      22 CAATGCTTTCTTGAACAAATATTCAAAACAAAAATTTCAAATTCAAAAATTTTCT 81

Qy     113 gctaaatcttaatttcccaaatttatctcaaaaaaaaaaaaaaaaaaaaaaaaaaa 172
           ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db      82 GCTAAATCTTTAATTTCCCAAATTTATCTCAAAAAAAAAAAAAAAAAAAAAAAAAAA 141

Qy     173 aaatgaaaaatgatgtaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 232
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142 AAATGAACAAGTACATGAAAAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 201

Qy     233 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 292
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     202 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 261

Qy     293 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     262 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 321

Qy     353 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     322 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 381
```

```
Qy     413 tttcttattctacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     382 TTTCTTTATTCTACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 441

Qy     473 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 532
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     442 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 501

Qy     533 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 592
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     502 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 561

Qy     593 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 652
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     562 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 621

Qy     653 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 712
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     622 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 681

Qy     713 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 772
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     682 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 741

Qy     773 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 832
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     742 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 801

Qy     833 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 892
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     802 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 861

Qy     893 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 952
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     862 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 921

Qy     953 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1012
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     922 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 981

Qy    1013 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1072
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     982 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1041

Qy    1073 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1042 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1101

Qy    1133 aaaaaaaaaaaaaaaaaa 1150
           ||||||||||||||||||
Db    1102 AAAAAAAAAAAAAAAAAA 1119
```

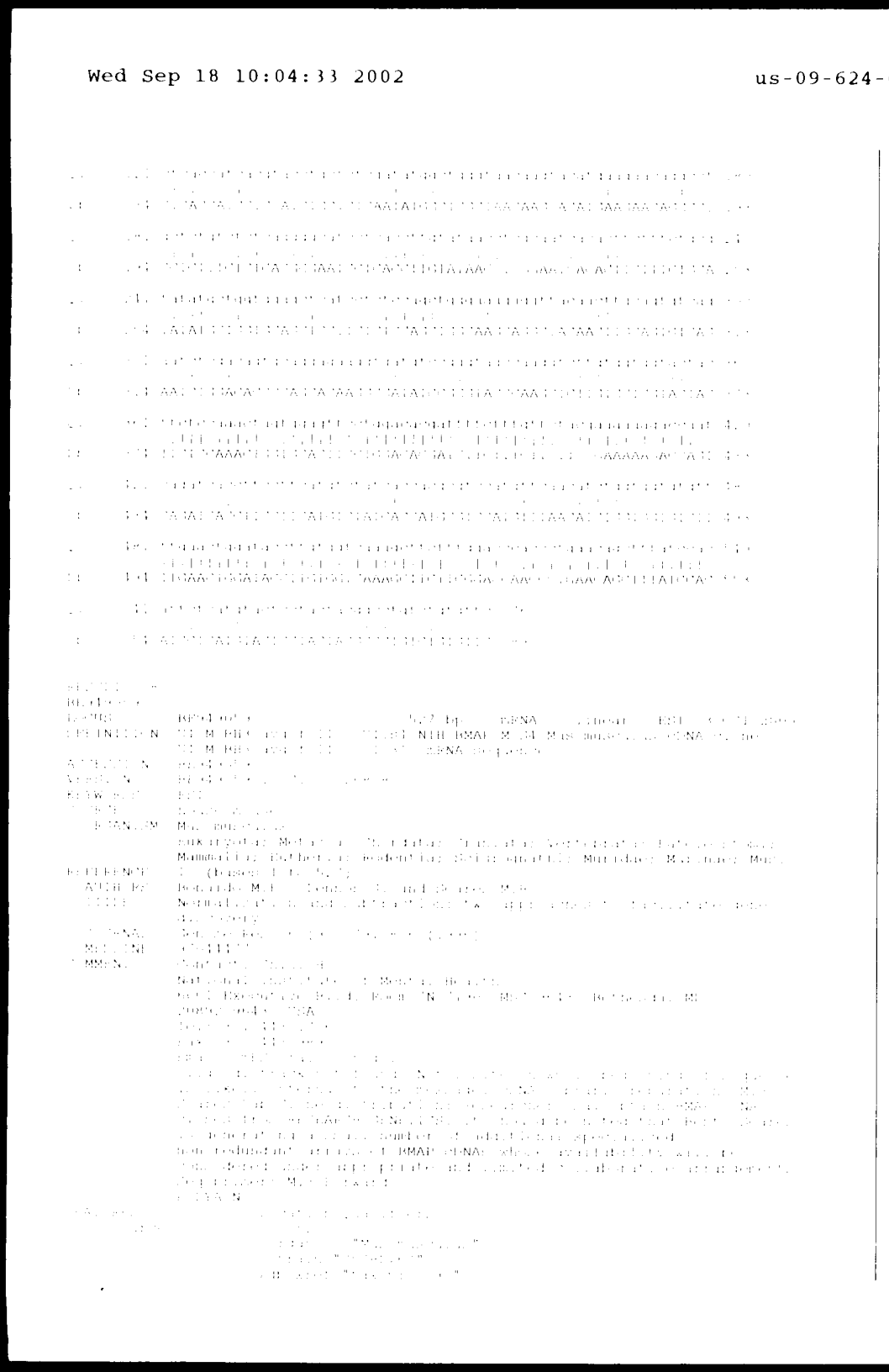
```
RESULT    3
BB654040
LOCUS       BB654040                661 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB654040 RIKEN full-length enriched, adult male liver tumor Mus
            musculus cDNA clone C730043M04 5', mRNA sequence.
ACCESSION   BB654040
VERSION     BB654040.1  GI:16487868
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 661)
  AUTHORS   Arakawa,T., Carninci,P., Sekida,Y., Furuno,M., Hasegawa,Y., Hara,A.
            , Hirakawa,M., Hori,F., Ishii,Y., Kitamura,K., Kawai,J., Kondo
            ,H., Koya,S., Matsuda,H., Miyazaki,A., Nakura,K., Ohno,M.,
            Sakazaki,Y., Saito,R., Sakai,K., Sakai,M., Sano,H., Susaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tozawa,A., Tsukahara,F., Wakeda,Y., Tanaka,T., Toyo.i.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse EST's (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp
            URL: http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1640 (2000)
```



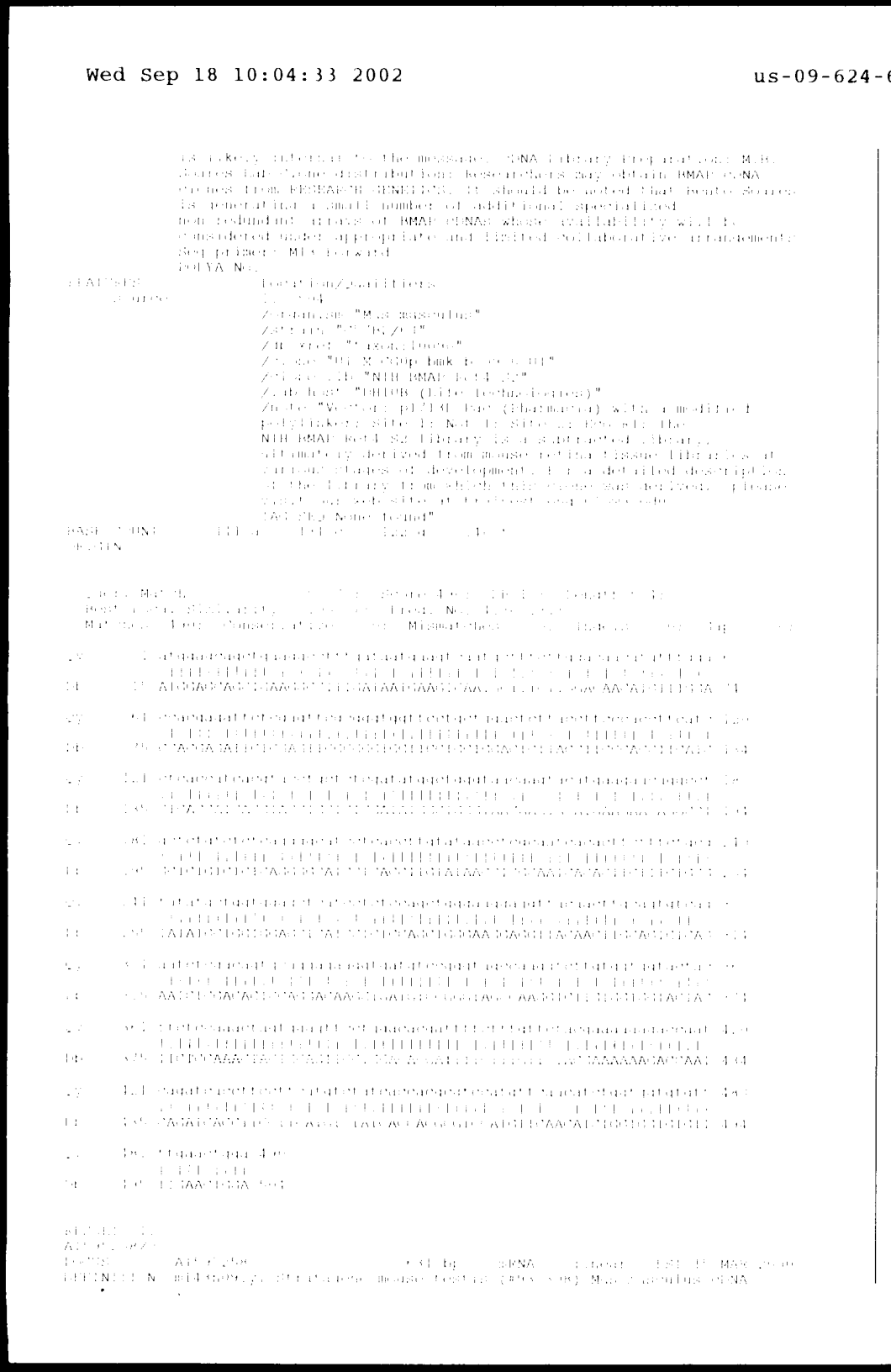




```
            prepared by using trehalose thermo activated reverse
            transcriptase and subsequently enriched for full length by
            cap trapper. cDNA went through one round of normalization
            to Rot = 20.1 and subtraction to Rot = 450.0. Second
            strand cDNA was prepared with the primer adapter of
            sequence [illegible]
            [illegible]. cDNA was cleaved with XhoI and BamHI. Vector: a
            modified pBluescript KS(+) after bulk excision from Lambda
            [illegible]. Retina RNA was provided by Stefano Gustincich,
            Department of Neurobiology, Harvard Medical School, 220
            Longwood Ave., Boston, MA 02115, USA, whose support we
            gratefully acknowledge."
BASE COUNT     [illegible] a   [illegible] c   [illegible] g   [illegible] t
ORIGIN

  Query Match             [illegible];  Score 544;  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative    [illegible];  Mismatches   [illegible];  Indels    4;  Gaps    4;

Qy          1 atggagcactgaagcctttgataatgaagtcaatgctttcttggacaacatgtttgga 60
Db         47 [illegible]

Qy         61 [illegible] 120
Db        107 [illegible]

Qy        121 [illegible] 180
Db        [illegible]

Qy        181 [illegible] 240
Db        [illegible] 286

Qy        241 [illegible]
Db        [illegible] 346

Qy        301 [illegible]
Db        [illegible] 406

Qy        361 [illegible] 420
Db        [illegible] 466

Qy        421 [illegible] 480
Db        467 [illegible] 526

Qy        481 [illegible] 538
Db        527 [illegible]

Qy        [illegible] 596
Db        [illegible] 646

Qy        [illegible] 620
Db        [illegible] 67[illegible]


RESULT [illegible]
AI[illegible]
LOCUS       AI[illegible]          589 bp    mRNA    linear   EST [illegible] JUL 1999
DEFINITION  UI-R-C2p-[illegible] UI-R-C2p Rattus norvegicus cDNA clone
            UI-R-C2p-[illegible] 3', mRNA sequence.
ACCESSION   AI[illegible]
VERSION     AI[illegible]  GI:[illegible]
KEYWORDS    EST.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 589)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (8), 791-806 (1996)
  MEDLINE   [illegible]
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: [illegible]
            Fax: [illegible]
            Email: msoares@blue.weeg.uiowa.edu
            Oligo-dT track not found, Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.
            Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
            through Research Genetics This clone is also available through the
            [illegible] IMAGE
            ID [illegible]
            Seq primer: M13 Forward
            POLYA N.
FEATURES             Location/Qualifiers
     source          1..589
                     /organism "Rattus norvegicus"
                     /strain "Sprague-Dawley"
                     /db_xref "taxon:10116"
                     /clone "UI-R-C2p-[illegible]"
                     /clone_lib "UI-R-C2p"
                     /dev_stage "adult"
                     /lab_host "DH10B (Life Technologies)"
                     /note "Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-C2p
                     library is a subtracted library derived from the UI-R-C1
                     library, which is a subtracted library derived from the
                     UI-R-C0 library. The UI-R-C0 library consisted of a
                     mixture of individually tagged normalized libraries
                     constructed from rat placenta, adult lung, brain, liver,
                     kidney, heart, spleen, ovary, muscle, 8, 12 and 18 day
                     embryo. The tag is a string of 3-5 nucleotides present
                     between the Not I site and the oligo-dT track which allows
                     identification of the library of origin of a clone within
                     the mixture. The subtracted library (UI-R-C2p) was
                     constructed as follows: PCR amplified cDNA inserts from
                     UI-R-C1 clones from which 3' ESTs had been derived was
                     used as a driver in a hybridization with the UI-R-C1
                     library in the form of single-stranded circles. The
                     remaining single-stranded circles (subtracted library) was
                     purified by hydroxyapatite column chromatography,
                     converted to double-stranded circles and electroporated
                     into E. coli bacteria (Life Technologies) to generate the
                     UI-R-C2p library. This procedure has been previously
                     described (Bonaldo, Lennon and Soares, Genome Research 6:
                     791-806, 1996)"
BASE COUNT     126 a    155 c    140 g    168 t
ORIGIN

  Query Match             [illegible];  Score 528;  DB [illegible];  Length 589;
  Best Local Similarity   94.8%;  Pred. No. 4.3e-142;
  Matches  [illegible];  Conservative    [illegible];  Mismatches   [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          1 [illegible] 60
Db         [illegible]

Qy         61 [illegible] 120
Db         74 [illegible]
```

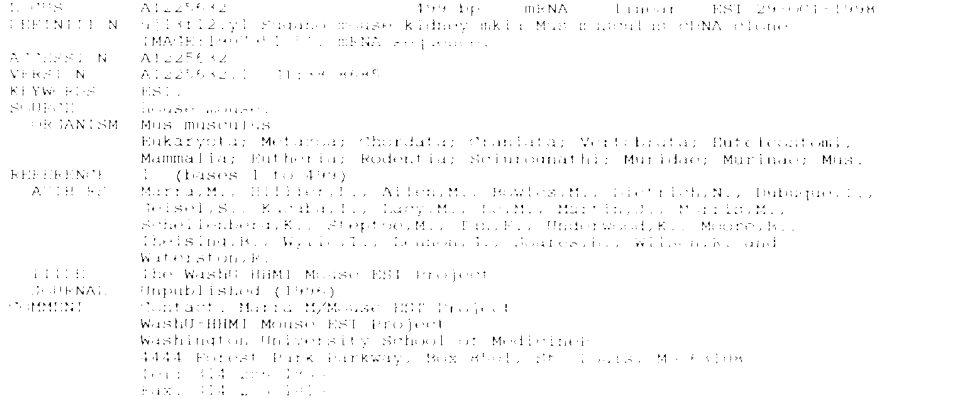


```
Db [illegible] [illegible]AA[illegible]AA [illegible]

RESULT [illegible]
AW4[illegible]
LOCUS       AW4[illegible]          521 bp    mRNA    linear   EST 24-FEB-2000
DEFINITION  UI-M-BH[illegible] NIH_BMAP_M_S4 Mus musculus cDNA clone
            UI-M-BH[illegible] 3', mRNA sequence.
ACCESSION   AW4[illegible]
VERSION     AW4[illegible]
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 521)
  AUTHORS   [illegible]
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N [illegible] MSC [illegible] Bethesda, MD
            20892-9643, USA
            Tel: [illegible]
            Fax: [illegible]
            Email: [illegible]
            [illegible] not found. Not I site shown in beginning of sequence
            is likely internal to the message. cDNA Library Preparation: M.B.
            Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
            clones from RESEARCH GENETICS. It should be noted that Bento Soares
            is generating a small number of additional specialized
            non-redundant arrays of BMAP cDNAs whose availability will be
            considered under appropriate and limited collaborative arrangements
            Seq primer: M13 Forward
            POLYA No.
FEATURES             Location/Qualifiers
     source          1..521
                     /organism="Mus musculus"
                     /[illegible]
                     /db_xref="taxon:10090"
                     /clone="UI-M-BH[illegible]"
                     /clone_lib="NIH_BMAP_M_S4"
                     /[illegible]
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_M_S4 library is a subtracted library of a series,
                     ultimately derived from a mixture of individually tagged
                     normalized libraries from ten regions of the mouse brain
                     (cerebellum, brain stems, olfactory bulbs, hypothalamus,
                     cortex, amygdala, basal ganglia, pineal gland, striatum,
                     hippocampus) after a series of subtractions to reduce the
                     representation of cDNAs from which ESTs had already been
                     generated. The following serially subtracted libraries
                     were generated in this process: NIH_BMAP_M_S4,
                     NIH_BMAP_M_S[illegible], NIH_BMAP_M_S[illegible], NIH_BMAP_M_S[illegible],
                     NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
                     (NIH_BMAP_M_S4) was constructed as follows. [illegible]
                     cDNA inserts from NIH_BMAP_M_S[illegible], NIH_BMAP_M_S[illegible] and
                     NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
                     was used as a driver in a hybridization with a pool of
                     the NIH_BMAP_M_S3.3, NIH_BMAP_M_S[illegible] and NIH_BMAP_M_S[illegible]
                     libraries in the form of single-stranded circles. The
                     remaining single-stranded circles (subtracted library)
                     was purified by hydroxyapatite column chromatography,
                     converted to double stranded circles and electroporated
                     into DH10B bacteria (Life-technologies) to generate the
                     NIH_BMAP_M_S4 library. This procedure has been previously
                     described (Bonaldo, Lennon and Soares, Genome Research
                     6:791-806, 1996)
                     [illegible] NIH_BMAP_M_S4
                     [illegible] brain stems
                     [illegible]
BASE COUNT      1[illegible] a    13[illegible] c    125 g    14[illegible] t    1 others
ORIGIN

  Query Match             5[illegible]%;  Score 49[illegible];  DB 9;  Length 521;
  Best Local Similarity   9[illegible]%;  Pred. No. [illegible]134;
  Matches  [illegible];  Conservative    ;  Mismatches    [illegible];  Indels    ;  Gaps    ;

Qy          1 [illegible] 60
Db         14 [illegible] 73

Qy         61 [illegible] 120
Db         74 [illegible] 133

Qy        121 [illegible] 180
Db        134 [illegible] 193

Qy        181 [illegible] 240
Db        194 [illegible] 253

Qy        241 [illegible] 300
Db        254 [illegible] 313

Qy        301 aat[illegible]caaggtcttgtggtggtactac 360
              [illegible]
Db        314 AATCTCGACAGTGCAGGAGAAGGTGATGTCCGGGTAGCCAAGG[illegible]GTGGTGGTACTAC 373

Qy        361 ttctccaaactagtggagttcctggacacgatttttctttgttctacgaaaaaagaccaat 420
              [illegible]
Db        374 [illegible] 433

Qy        421 [illegible] 480
Db        434 [illegible] 493

Qy        481 [illegible] 504
Db        494 [illegible]AA 517


RESULT 10
BF461032
LOCUS       BF461032                504 bp    mRNA    linear   EST 04-DEC-2000
DEFINITION  UI-M-CG0p-bmk-b-09-0-UI.s1 NIH_BMAP_Rc14_S2 Mus musculus cDNA clone
            UI-M-CG0p-bmk-b-09-0-UI 3', mRNA sequence.
ACCESSION   BF461032
VERSION     BF461032.1  GI:[illegible]
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 504)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N 7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1764
            Fax: 301 443 9890
            Email: bmEST@mail.nih.gov
            [illegible] not found. Not I site shown in beginning of sequence
```

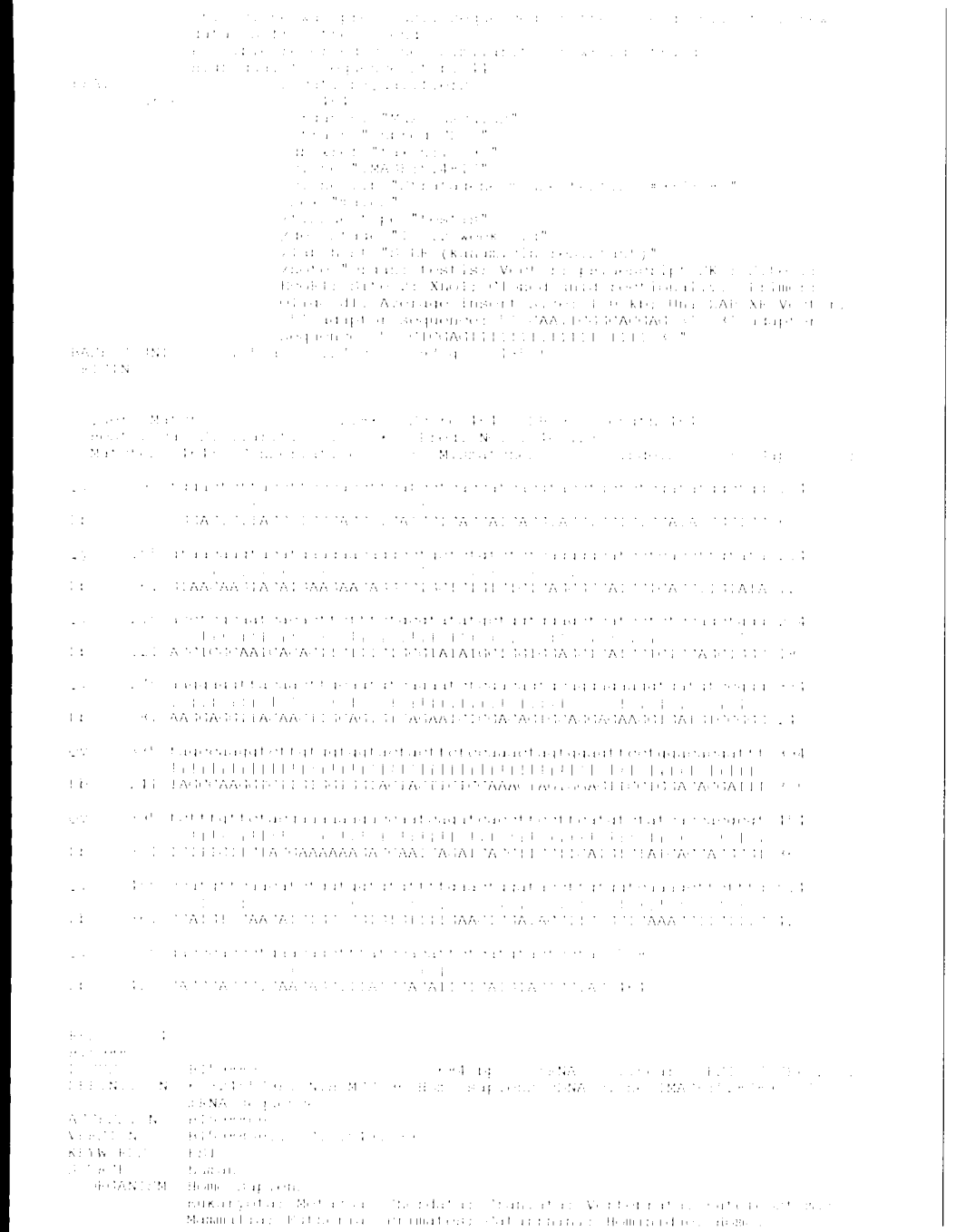


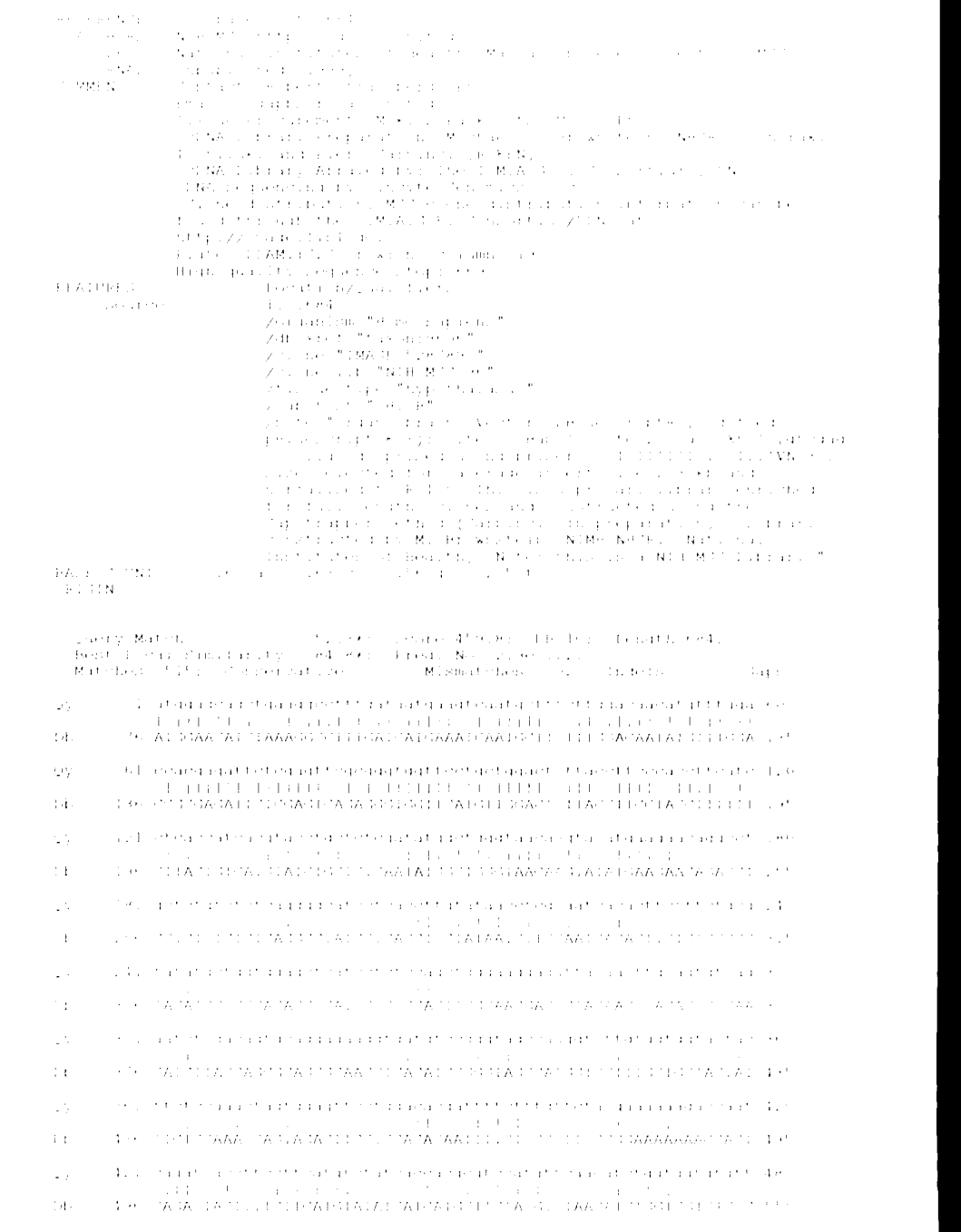
RESULT 12
AI228562
LOCUS AI228562 709 bp mRNA linear EST 25-OCT-1998
DEFINITION ml1d12.y1 Stratagene mouse kidney #937120 Mus musculus cDNA clone IMAGE:1400742 5', mRNA sequence.
ACCESSION AI228562
VERSION AI228562.1 GI:3144666
KEYWORDS EST.
SOURCE mouse, mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 709)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Gibbs,R., Dubuque,T., Delehaunty,K., Kalicki,J., Kucaba,T., Martin,J., Miller,N., Schellenberg,K., Stoptkin,M., Underwood,K., Moore,B., Carlston,B., Wylie,T., Leonard,S., Bonfield,J., Wilcox,L. and Waterston,R.
TITLE The WashU HHMI Mouse EST Project
JOURNAL Unpublished (1998)
COMMENT Contact: Marco M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
1444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
M81:976147
Seq primer: custom primer used
High quality sequence stops: 470.
FEATURES Location/Qualifiers
source 1..709
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1400742"
/clone_lib="Stratagene mouse kidney #937120"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="vector: pBluescript II SK(-); Site_1: BrainI (CATCGATG); site_2: BrainI (CATCGATG); EST sequence from 3' primer M13 (-21); cDNA was primed with oligo(dT) containing a BrainI site; double stranded cDNA was ligated to a BrainI adaptor (5'-TCGACCCACGCGTCCG-3'), digested with NotI and BrainI; BrainI sites at the pBluescript polylinker (5' site=CATCGATG, 3' site=CATCGATG), xhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumi Surzo (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end primer CGACCTCGAGTCGACGCA."
BASE COUNT 186 a 137 c 218 g 164 t
ORIGIN

Query Match 52.8%; Score 464.2; DB 9; Length 499;
Best Local Similarity 99.4%; Pred. No. 1.6e-123;
Matches 466; Conservative 0; Mismatches 3; Indels 0; Gaps 0;



RESULT 13
AI428130
LOCUS AI428130 456 bp mRNA linear EST 07-MAR-1999
DEFINITION mi13f07.y1 Stratagene mouse testis (#937208) Mus musculus cDNA clone IMAGE:1414817 5', similar to TR:Q20308 Q20308 SIMILAR TO S. CEREVISIAE INTEGRAL MEMBRANE PROTEIN YNL1..., mRNA sequence.
ACCESSION AI428130
VERSION AI428130.1 GI:4274056
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 464)
AUTHORS Marra,M., Hillier,L., Kucaba,T., Martin,J., Bowles,M., Wylie,T., Waterston,R. and Wilson,R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marco Marra M/Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, Missouri, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.





```
Qy      481 ttgaactagataccttgtgatcaaaactttctttgaccccacctcgaacaagctttatccac 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      556 TTGAACTGGATATCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCAC 615

Qy      541 attctcatatgtactcctactacgcctgtctgtgtgtccatccatgcacaaaatacctttgg 600
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      616 ATTCTTATCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC 675

Qy      601 taaaaa 607
            ||||||
Db      676 GGGAAGA 682


RESULT 15
BF470691/c
LOCUS       BF470691                461 bp    mRNA    linear   EST 04-DEC-2000
DEFINITION  UI-M-BH3-avq-t-11-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
            UI-M-BH3-avq-t-11-0-UI.r1, mRNA sequence.
ACCESSION   BF470691
VERSION     BF470691.1  GI:11549824
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 461)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd, Room 7149, MSC 9641, Bethesda, MD
            20892 9641, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
            should be noted that Bento Soares is generating a small number of
            additional specialized non-redundant arrays of BMAP cDNAs whose
            availability will be considered under appropriate and limited
            collaborative arrangements
            Seq primer: M13 Reverse.
FEATURES             Location/Qualifiers
     source          1..461
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-BH3-avq-t-11-0-UI"
                     /clone_lib="NIH_BMAP_M_S4"
                     /dev_stage="27-32 days"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_M_S4 library is a subtracted library of a series,
                     ultimately derived from a mixture of individually tagged
                     normalized libraries from ten regions of the mouse brain
                     (cerebellum, brain stems, olfactory bulbs, hypothalamus,
                     cortex, amygdala, basal ganglia, pineal gland, striatum,
                     hippocampus) after a series of subtractions to reduce the
                     representation of cDNAs from which ESTs had already been
                     generated. The following serially subtracted libraries
                     were generated in this process: NIH_BMAP_M_S4,
                     NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
                     NIH_BMAP_M_S2, NIH_BMAP_M_S1.  The subtracted library
                     (NIH_BMAP_M_S4) was constructed as follows:  PCR-amplified
                     cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
                     NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
                     was used as a driver in a hybridization with a pool of
                     the NIH_BMAP_M_S2.1, NIH_BMAP_M_S1.2, and NIH_BMAP_M_S1.1
                     libraries in the form of single-stranded circles.  The
                     remaining single-stranded circles (subtracted library)
                     was purified by hydroxyapatite column chromatography,
                     converted to double-stranded circles and electroporated
                     into DH10B bacteria (LifeTechnologies) to generate the
                     NIH_BMAP_M_S4 library.  This procedure has been previously
                     described (Bonaldo, Lennon and Soares, Genome Research
                     6:791-806, 1996)"
BASE COUNT      127 a    108 c    121 g    104 t      1 others
ORIGIN


  Query Match          51.5%;  Score 452.4;  DB 10;  Length 461;
  Best Local Similarity   99.6%;  Pred. No. 4.4e-120;
  Matches  453;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 atgaagcaactctaaagcctttgataatgaaatcaatgctttttcttgaacaacatgtttgaa 60
            ||||||||||||||||||||||||||||||||| ||||||| |||||||||||||||||||| ||||
Db      455 ATCGAGCAGCTCAAGGCCTTTCATAATGAACTCAATGCTTTCTTGGACAACATGTNTGGA 396

Qy       61 ccacgaaatttctcgaatttcgcgaatgagttcctgtcgaactttacctttccaattttatc 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db      395 CCACGAGATTCTCGAGTTCGCGCTGGTTCCTGCTCGACTCTTACCTCCCACCTTCATC 336

Qy      121 ctcaccatcacgtacctgctctctatatgctgagtaacaagtacatgaagaacagggct 180
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      335 CTCACCATCACGTACCTGCTCTCGATATGGCTGGGTAACAAGTACATGAAGAACAGGGCT 276

Qy      181 gctctgtctctctcaggggcatcctcaccttgtataacctccgcaatcacacttctttctcgcg 240
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      275 GCTCTGTCTCTCAGGGGCATCCTCACCTTGTATAACCTCGGCAATCACACTTCTTTCTCTGG 216

Qy      241 tatatgctgtgtgaagctcatcctctccagctgtgaagtgagttgccaaacttgccagtgtgag 300
            |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db      215 TATATGCTGGTGTGAGCTCATCCTCTGCAGCTGGGAAGGAGTTACAACTTGCAGTGTGCAG 156

Qy      301 aatctcgacagtgcgaaaaaaaagtatatcgggtatgccaaagtctttgtgtgtgctacgc 360
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db      155 AATCTCGACAGTGCAGGAGAAGGTGATGTCGGGTAGCCAAGGTCTTGTGGTGCTACTAC 96

Qy      361 ttctccaaactaatgagttcctgcgacgcgatttctttgttgttctacgaaaaaagacgcaat 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       95 TTCTCCAAACTAATGAGTTCCTGCGACGCGATTTCTTTGTTGTTCTACGAAAAAAGACGCAAT 36

Qy      421 caaatcacttttcttcatgtctatgcaccacgcgtc 455
            |||||||||||||||||||||||||||||||||||
Db       35 CAGATCAGCTTCTTCATGTCTATGCACCACGCGTC 1


Search completed: September 17, 2002, 18:00:08
Job time: 5803 secs
```

Software version 4.0
Copyright (c) 1993 - 2002 Compugen Ltd.

GenCore version 5.1.6 Nucleic - nucleic search, using sw model

Run on: September 17, 2002, 16:26:05 ; Search time 7.12 Seconds
(without alignments)
4774.174 Million cell updates/sec

Title:
Perfect score: 4774
Sequence: 1 gatttttcatttttttttt.....ttttttttttttttttttt 4774

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3865688 seqs, 12281675? residues

Total number of hits satisfying chosen parameters: 7731376

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/backfiles1.seq:*
5: /cgn2_6/ptodata/2/ina/backfiles2.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles3.seq:*

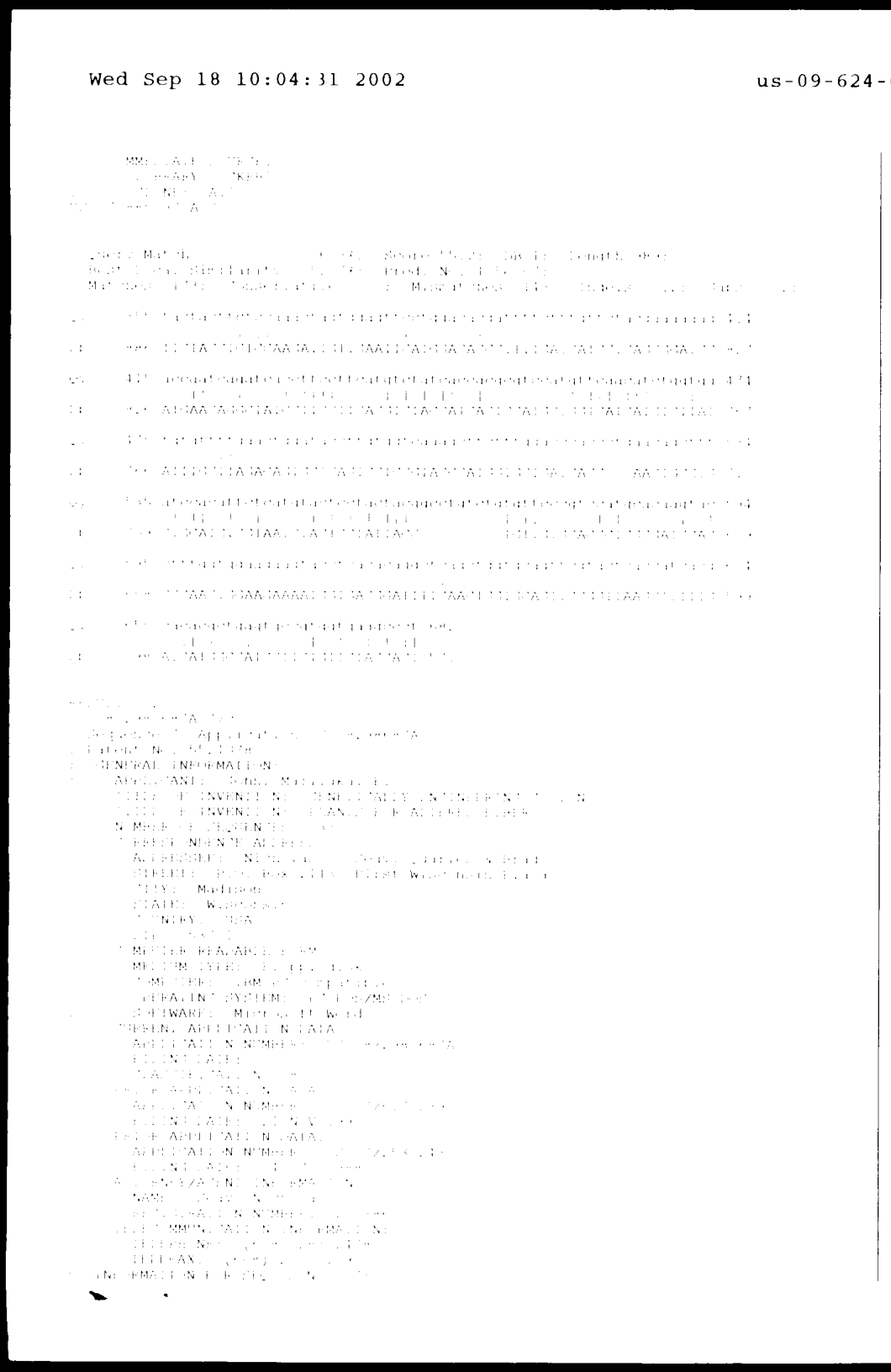
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

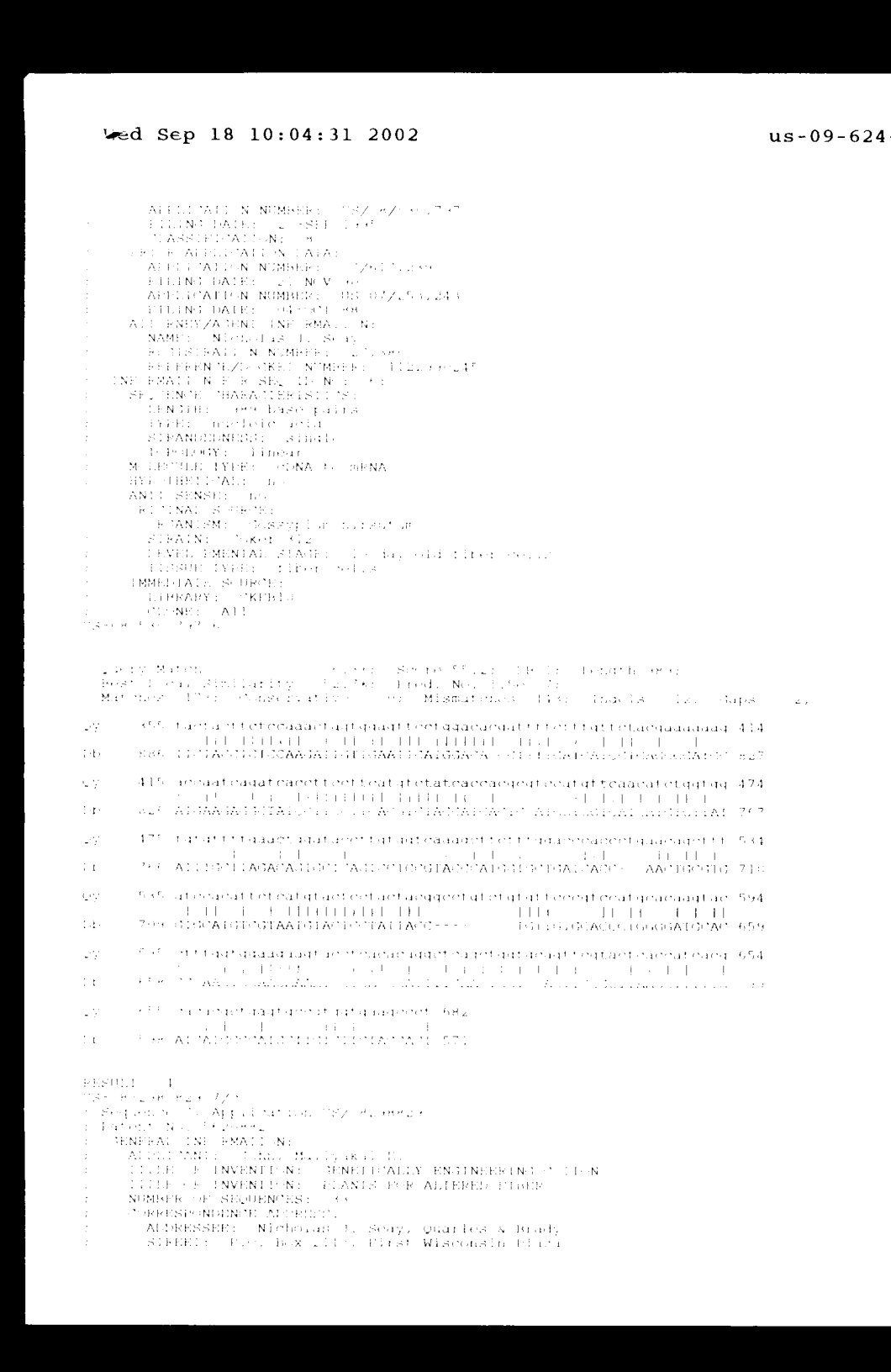
SUMMARIES

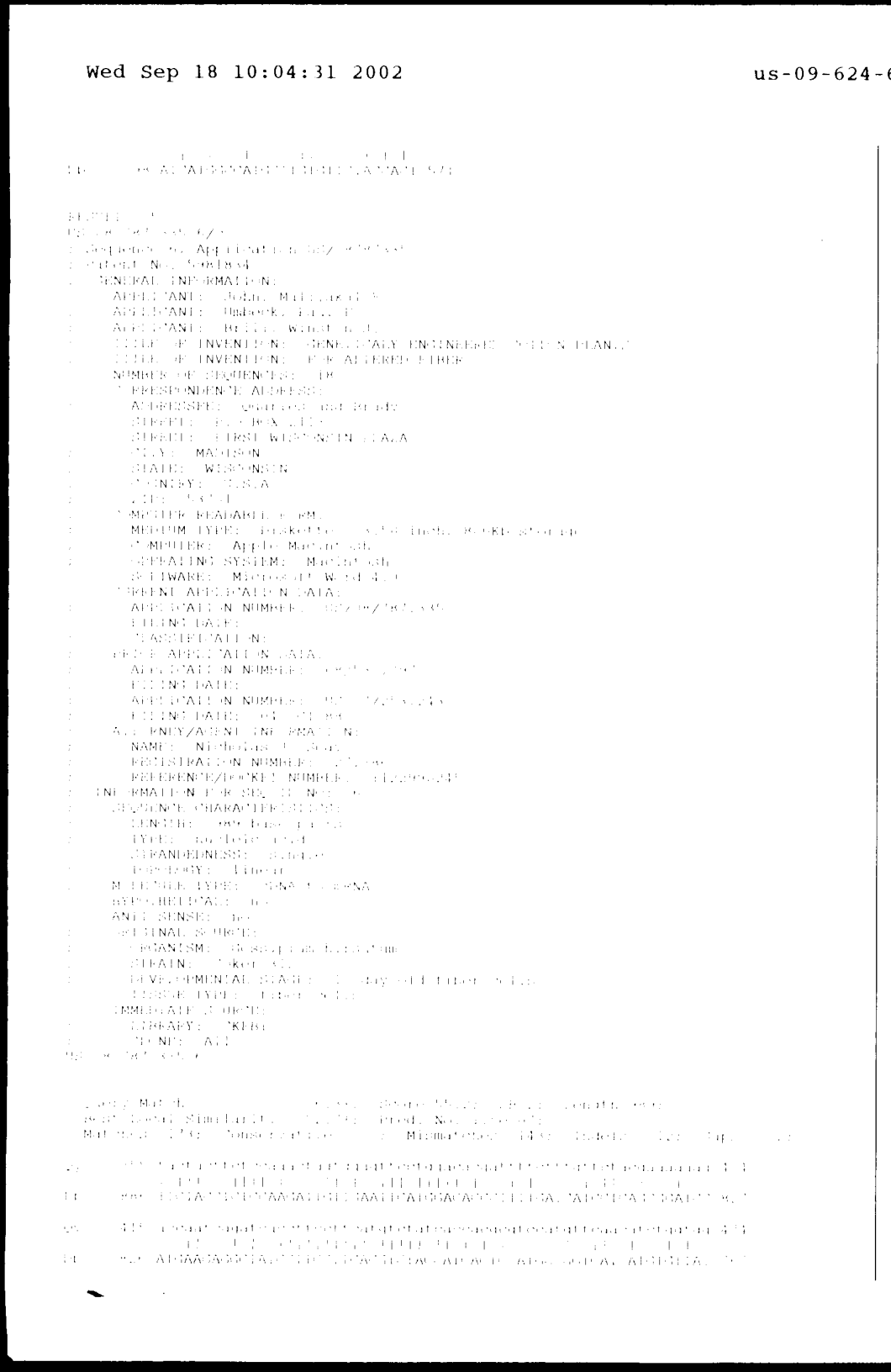
Result No.	Score	Query Match	Length	DB ID	Description
1	4774	100.0	4774	1	US-07-805-967-A
2	4774	100.0	4774	1	US-08-298-687-A
3	4774	100.0	4774	2	US-08-840-747-A
4	4774	100.0	4774	2	US-08-298-687-A
5	4774	100.0	4774	2	US-08-840-747-A
6	71	1.5	71	5	US-08-232-164-13
7	57.4	1.2	24	4	US-09-172-708-1F
8	50.2	1.1	201	4	US-09-007-[illegible]
9	50.2	1.1	201	4	US-09-244-790-1F
10	44	0.9	6496	4	US-08-822-962-12
11	44	0.9	6556	1	US-08-847-645-5
12	44	0.9	6556	5	US-08-847-645-7
13	44	0.9	6556	1	US-08-827-962-16
14	44	0.9	6556	4	US-08-827-962-2
15	43.6	0.9	5425	1	US-08-814-[illegible]
16	43.6	0.9	5425	1	US-08-814-[illegible]
17	43.6	0.9	5425	1	US-08-847-[illegible]
18	43.6	0.9	5425	1	US-08-[illegible]
19	43.6	0.9	5425	4	US-08-884-[illegible]
20	43.4	0.9	6556	1	US-08-861-[illegible]
21	42.6	0.9	6556	4	US-09-227-457-[illegible]
22	42.6	0.9	1468	4	US-08-972-927-8
23	42.4	0.9	5515	4	US-09-498-193-98
24	42.2	0.9	4095	6	5231168-1
25	42	0.9	5481	1	US-07-707-467-1
26	41.6	0.9	5556	1	US-08-847-[illegible]
27	41.6	0.9	5556	1	US-08-847-[illegible]
28	41.6	0.9	5556	1	US-08-467-781-1
29	41.6	0.9	5556	1	US-08-146-487-1
30	41.6	0.9	5556	2	US-08-483-924-5
31	41.6	0.9	5556	5	PCT-US91-06150-1
32	40.8	0.9	1189	1	US-07-781-014-4
33	40.8	0.9	1189	5	PCT-US92-09328-4
34	40.4	0.8	1373	5	PCT-US92-02677-6
35	40.4	0.8	1373	5	PCT-US95-04632-4
36	40.4	0.8	2028	2	US-08-983-750C-73
37	40.4	0.8	2028	5	US-09-234-613-73
38	40.2	0.8	261	2	US-08-774-045A-4
39	40.2	0.8	261	4	US-09-244-843-4
40	40.2	0.8	5253	2	US-08-256-710-3
41	40.2	0.8	8735	5	US-08-782-891-79
42	40	0.8	462	4	US-09-328-111-799
43	40	0.8	1441	1	US-07-814-994-9
44	40	0.8	1444	1	US-08-258-442-9
45	40	0.8	1444	1	US-08-328-809-4

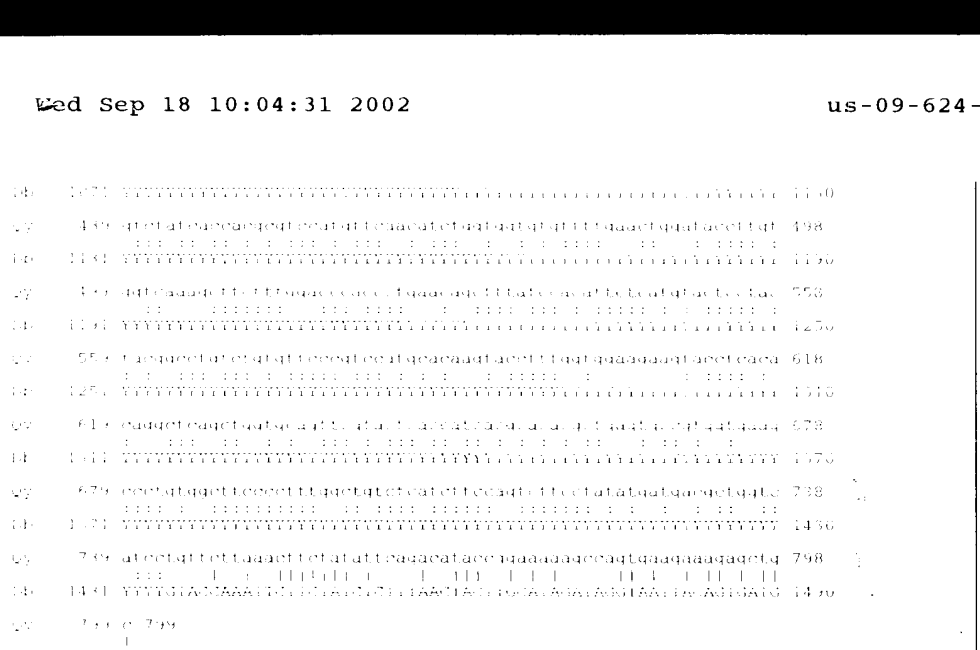
ALIGNMENTS

RESULT 1
US-07-805-967-A
; Sequence 1, Application US/07805967A
; Patent No. 5481168
; GENERAL INFORMATION:
; APPLICANT: Ellis, Marlyn H.
; TITLE OF INVENTION: GENETICALLY ENGINEERING FIBER
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nicholas J. Godici & Associates
; STREET: 1 [illegible] First Wisconsin Plaza
; CITY: Madison
; STATE: Wisconsin
; COUNTRY: USA
; ZIP: 53711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/805,967A
; FILING DATE: 10-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/[illegible]
; FILING DATE: 21-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/[illegible]
; FILING DATE: 14-JUL-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Godici, Nicholas J.
; REGISTRATION NUMBER: 27,896
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 283-2470
; TELEFAX: (608) 257-3719
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4774 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Gossypium hirsutum
; STRAIN: Coker 312
; DEVELOPMENTAL STAGE: 12 day old fiber cells
; TISSUE TYPE: fiber cells







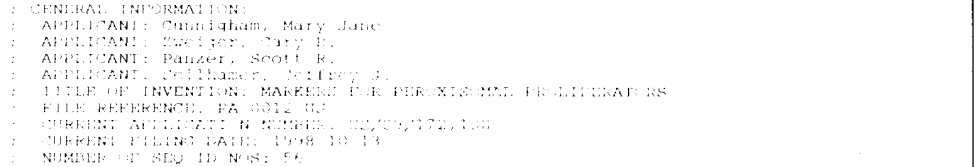


RESULT 7
US-09-172-108-45
; Sequence 45, Application US/09172108
; Patent No. 6160104
; GENERAL INFORMATION:
; APPLICANT: Cunningham, Mary Jane
; APPLICANT: Spellen, Cary E.
; APPLICANT: Panzer, Scott R.
; APPLICANT: Gellhaus, Jeffrey S.
; TITLE OF INVENTION: MARKERS FOR PERIODONTAL PATHOGENS
; FILE REFERENCE: PA 0012 US
; CURRENT APPLICATION NUMBER: US/09/172,108
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PERL Program
; SEQ ID NO 45
; LENGTH: 124
; TYPE: DNA
; ORGANISM: Bacteroides
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 34, 53, 87, 109, 122
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; OTHER INFORMATION: 70515905H1
US-09-172-108-45

Query Match 5.1%; Score 47.4; DB 4; Length 124;
Best Local Similarity 65.6%; Pred. No. 1.2e-05;
Matches 80; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

RESULT 8
US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Socransky, Sigmund S.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Ribo
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/450003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-05
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match 4.6%; Score 40.2; DB 4; Length 289;
Best Local Similarity 8.8%; Pred. No. 0.0047;
Matches 13; Conservative 87; Mismatches 100; Indels 0; Gaps 0;



RESULT 9
US-09-244-796-17/c
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Socransky, Sigmund S.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Ribo
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/250007
; CURRENT APPLICATION NUMBER: US/09/244,796
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-05

RESULT 13
US-08-624-670-7
; Sequence 7, Application US/08624670
; Patent No. 6258944
; GENERAL INFORMATION:
; APPLICANT: B.SENBLOM, MARIES, L.
; APPLICANT: VAN DER PL.ST, LE NAEVET, D.L.
; APPLICANT: GRESSER, SAMUEL A.
; APPLICANT: TULLY, PETER, F.
; APPLICANT: HERR, JOHN W.
; APPLICANT: DITA, MICHAEL, B.
; APPLICANT: CHUN, JANY
; TITLE OF INVENTION: TILLIN ASSAY
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JOANNE M. GIESSER MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE P.O. BOX 2000
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/624,670
; FILING DATE: 21-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,851
; FILING DATE: 02-APR-1996
; APPLICATION NUMBER: US/08/331,002
; FILING DATE: 15-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: GIESSER, JOANNE M
; REGISTRATION NUMBER: 32,933
; REFERENCE/DOCKET NUMBER: 19642Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 732 594 4046
; TELEFAX: 732 594 1720
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 456 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-624-670-7

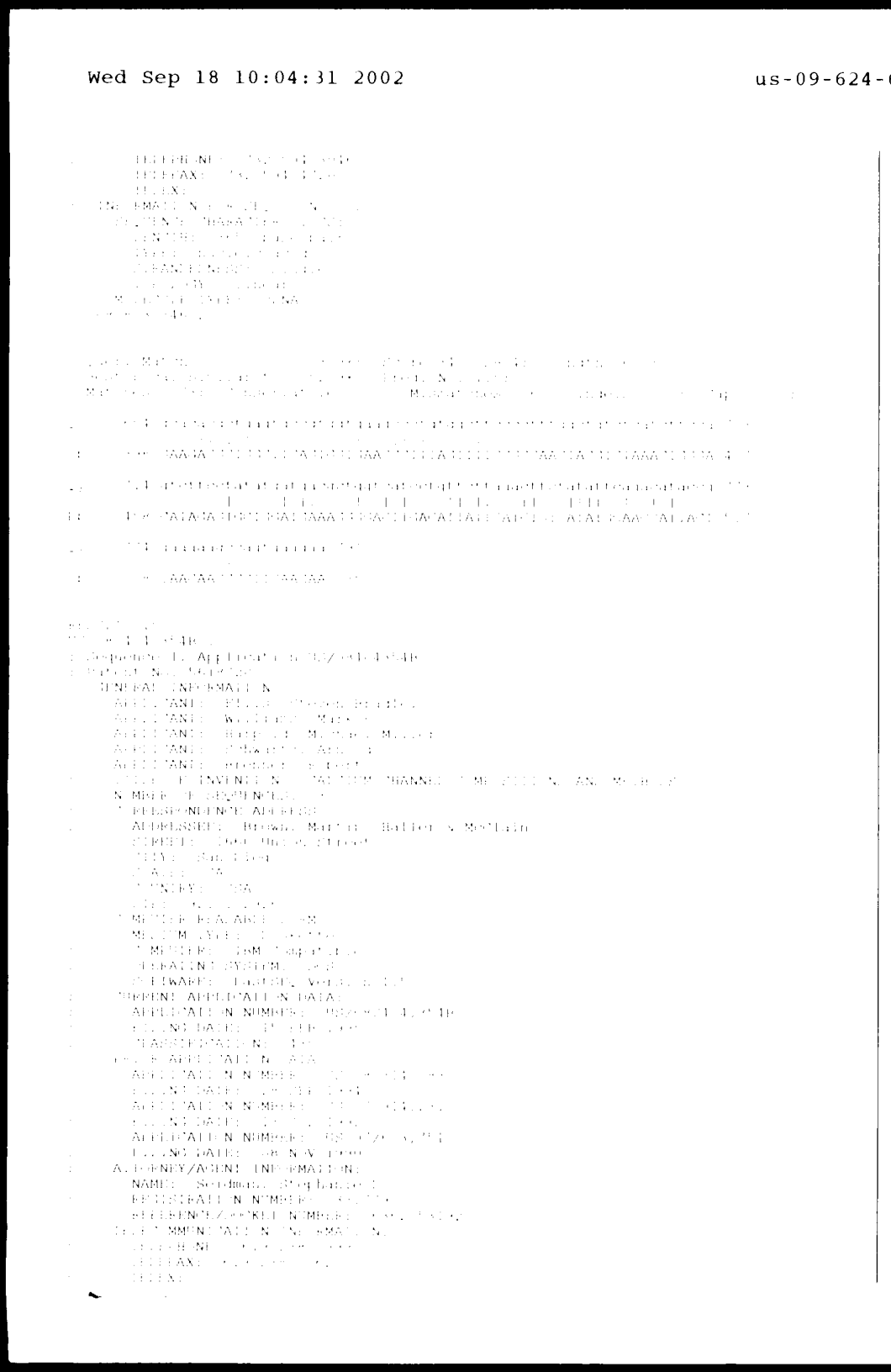
Query Match 5.9%; Score 34; DB 4; Length 456;
Best Local Similarity 52.9%; Pred. No. 1.5;
Matches 72; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

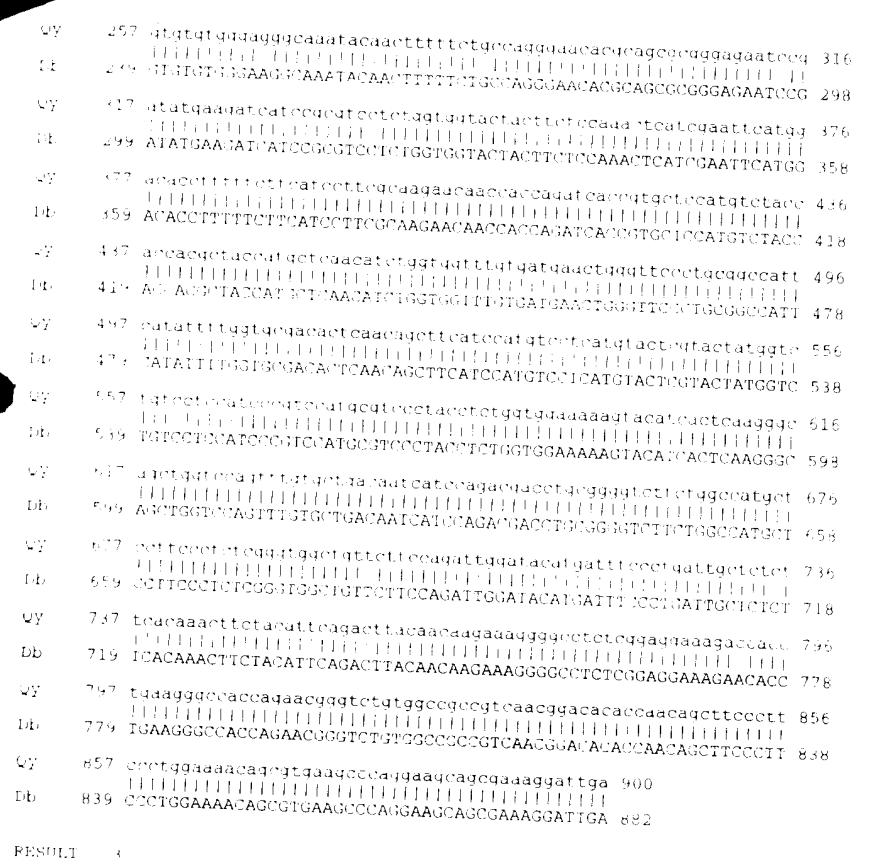
Qy 654 gaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 713
Db 398 gaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 457

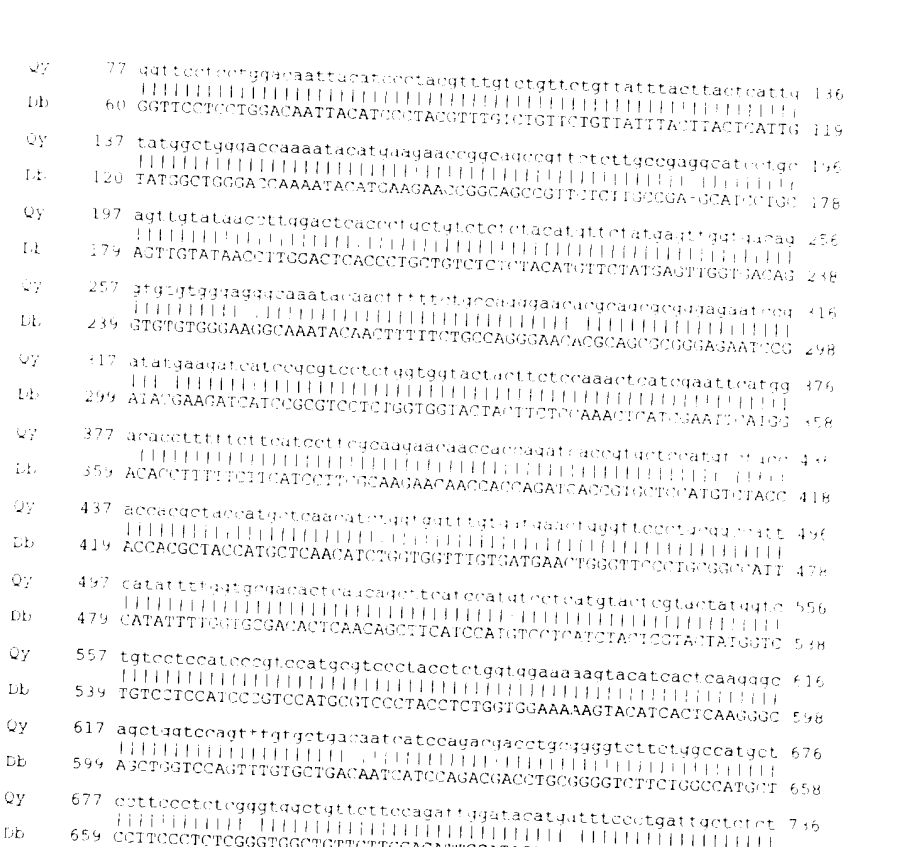
Qy 714 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 773
Db 458 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 517

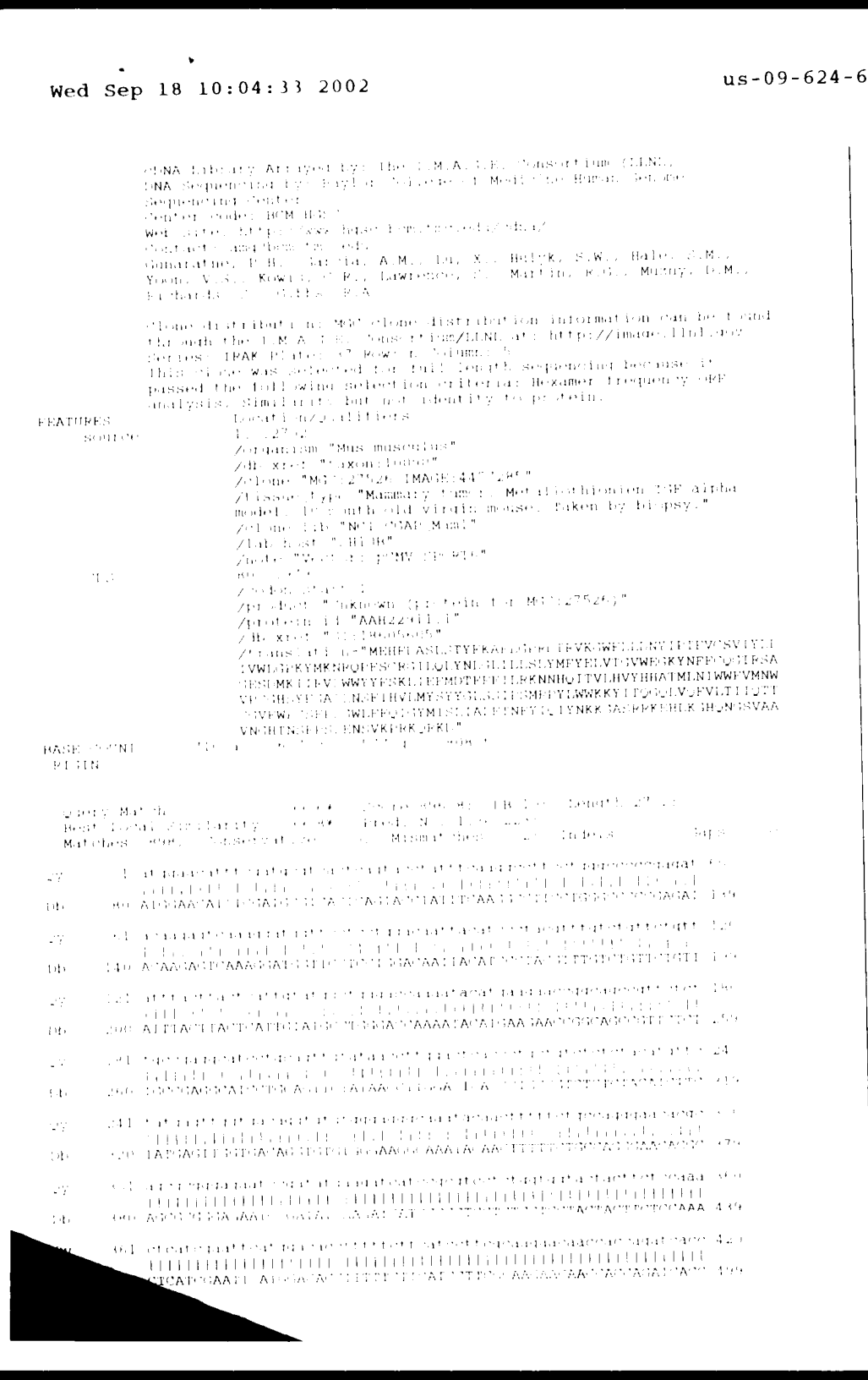
Qy 774 aaaaaaaaaaaaaaaaa 791
Db 518 aaaaaaaaaaaaaaaaa 535

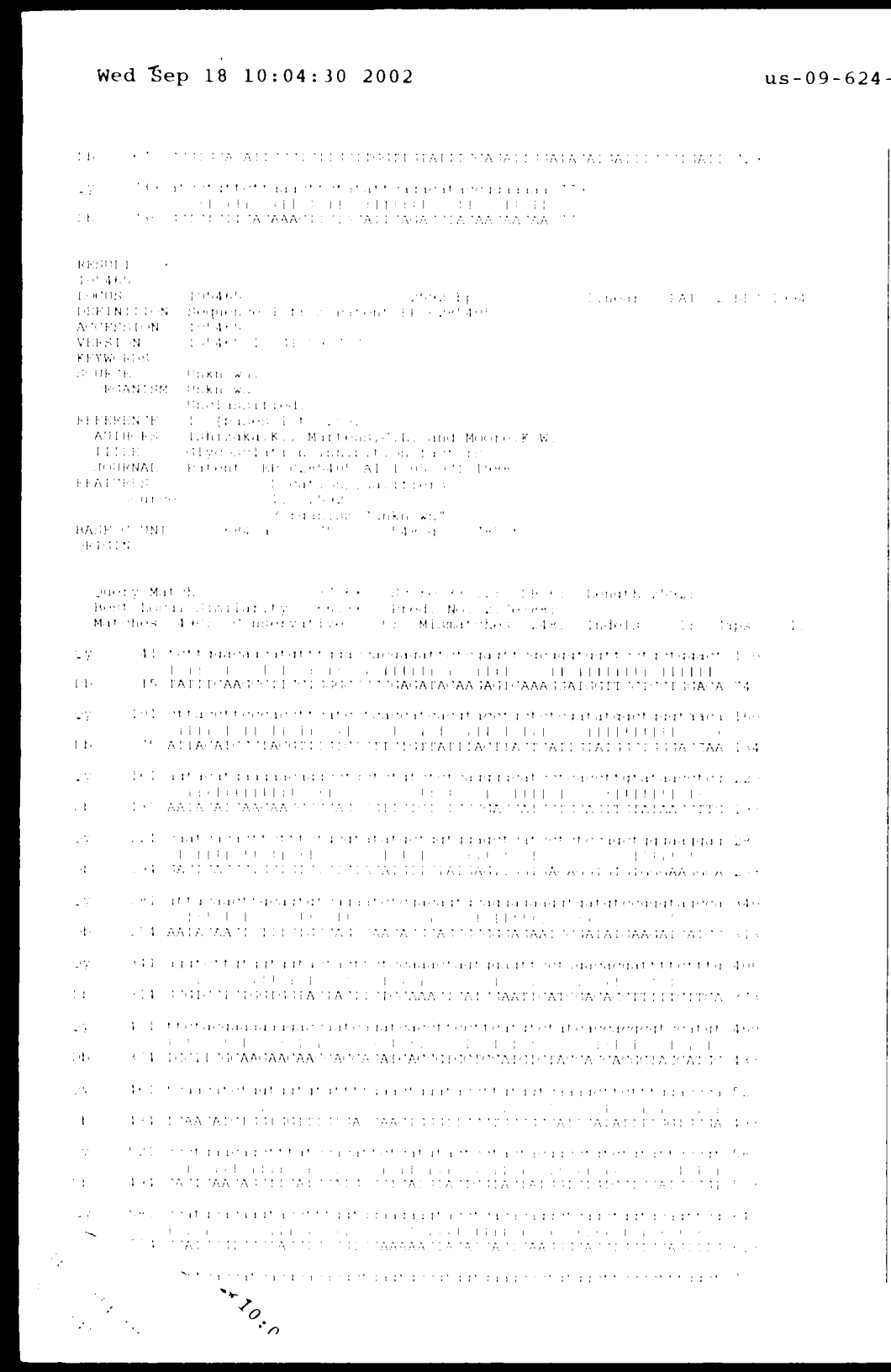
RESULT 14
US-08-624-670-2
; Sequence 2, Application US/08624670
; Patent No. 6258944
; GENERAL INFORMATION:
; APPLICANT: HERR, JOHN W.
; APPLICANT: CASKEY, C. THOMAS
; APPLICANT: CHU, GONGYUN
; APPLICANT: PHILLIPS, MICHAEL SEAN
; TITLE OF INVENTION: RAT OR RECEPTORS AND NUCLEIC ACIDS
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JOANNE M. GIESSER MERCK & CO., INC.
; STREET: 126 EAST LINCOLN AVENUE P.O. BOX 2000
; CITY: RAHWAY
; STATE: NJ
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/624,670
; FILING DATE: 21-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: GIESSER, JOANNE M
; REGISTRATION NUMBER: 32,933
; REFERENCE/DOCKET NUMBER: 19642Y
; TELECOMMUNICATION INFORMATION:











```
  JOURNAL   Submitted (21-JUL-1999) The Arrhenius Laboratories F3, The
            Wenner-Gren Institute, Stockholm University, Stockholm S-106 91,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .3709
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /sex="male"
                     /tissue_type="liver"
                     /dev_stage="9.11 weeks"
     gene            1. .3709
                     /gene "Ssc2"
     CDS             37. .915
                     /gene="Ssc2"
                     /function="putatively involved in fatty acid elongation"
                     /note="membrane protein"
                     /codon_start=1
                     /product="SSC2"
                     /protein_id="AAF72573.1"
                     /db_xref="GI:9101521"
                     /translation="MEQLKAFLNGVNAFLQNMFGPFISVVHGWFLLGSYLPTFILIII
                     YLLSIWLCNKYMKNPPALSLPGILTLYNLAITLLSAYMLVELILFSWEGGYNLQCQNI
                     DSAGPGDVRVAKVLWWYYFSKIVEFLDTIFFVLRKKTNQITFLHVYHHACMPNIWWCV
                     LNWIPCGQSFFGPTLNSFIHIIMYSYYGLSVFPSMRPYLWWKKYLTQAQLVQFVLTIT
                     HTISAVYVPCDFPCSIPNSAVMGMIMVTVLFLFSNFYVQTYIKPSDKQKKSK
                     KAHLIVANGMTDKKAQ"
BASE COUNT       894 a    867 c    774 g    1079 t
ORIGIN


  Query Match             100.0%;  Score 879;  DB 12;  Length 879;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-254;
  Matches  879;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 atggaggcagctgaaggcctttgatagtgaagtgaatgctttcttggacaacatagttgtgca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37 ATGGAGGCAGCTGAAGGCCTTTGATAATGAAGTCAATGCTTTCTTGGACAACATAGTTGTGA 96

Qy      61 ccaagagattctcgagttcgagaggtagttcctactgaactcttacctcccacctttcatc 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      97 CCACGAGATTCTCGAGTTCGCGGGTGGTTCCTGCTGGACTCTTACCTTCCCACCTTCATC 156

Qy     121 ctcaccatcacgtacctgttctcgatatgactgctgtaacaagtacatgcaagaacaggcct 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     157 CTCACCATCACGTACCTGTTCTCGATATGACTGCTGTAACAAGTACATGCAAGAACAGGCCT 216

Qy     181 gctctgtctctcagagggcatcctcaccttgtataacctcgcaatcacactttttgtgcag 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     217 GCTCTGTCTCTCAGAGGGCATCCTCACCTTGTATAACCTCGCAATCACACTTTTTGTGCAG 276

Qy     241 tatatgctggtggaggctgtacctctccagctggagaaagaggttacaactttgcagtatcag 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     277 TATATGCTGGTGGAGCTCATCCTCTCCAGCTGGGAAAGAGGTTACAACTTTGCAGTGTCAG 336

Qy     301 aatctcgacaagtgcaagaagaaaatcaatgtcgagtacgcaagctcttgtgattgatgact 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     337 AATCTCGACAGTGCAGGAGAAGCTCAATGTCCGCGTACGCAAGCTCTTGTGCTGCTGACTAC 396

Qy     361 ttttccaaactagtggaattcttgaacacgatttttctttgtttctacgaaaaaaaaccaat 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     397 TTCTCCAAACTAGTGGAGTTCCTGAACACGATTTTTCTTTGTTTCTACGAAAAAAGACCAAT 456

Qy     421 cgaatcgacttcttcatgatgtatgtcgcgctcatgttcaacatctggtggtgtgttt 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     457 CGAATCACCTTCCTTCATGTCTATCACCACGCGTCCATGTTCAACATCTGGTGGTGTGTT 516

Qy     481 ttgaactggatacctttgtggtcaaagcttctttggacccaccctgaacagctttatccac 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     517 TTGAACTGGATACCTTTGTGGTCAAAGCTTCTTTGGACCCACCCTGAACAGCTTTATCCAC 576

Qy     541 attctcatgtactctactagcgctgtctgtgttcagtccatgcaaaagtacctttgg 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     577 ATTCTCATGTACTCTACTAGCGCTGTCTGTGTTCAGTCCATGCAAAAGTACCTTTGG 634

Qy     601 tggaaagaagtaactcctcaagcagctgaactgatgcagttcctactcaccatcatcgtcgct 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     635 TGGAAGAAGTACCTCACACAGCCTCAGTTGGTGCTGTTCGTACTCACCATCATCGTCGCT 694

Qy     661 ctgaatgccctgatgtgaggcctatgcttcctttgctgtctctcatttcgatgtttcctgt 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     695 CTGAGTGCCGTGGTGAAGGCCTCTGGCTTCCCCTTGCTGTCTATTTTCAGTTTCCTGTT 754

Qy     721 tatatgatgcagcgtgatcatcttttgaagcttctatattcgaagtatgcttcaaaaaaa 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     755 TATATGATGAGCGTGATCATCTTCTTGAAGCTTCTATATTCGAAGTATGCTTCAAAAAAA 814

Qy     781 cgaatgaaaaaaaaaatgaatgcaaaaaaaaaaaaatgaaaaaatgatttcttcgaaaaaa 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     815 CGAATGAAAAAAAAATGAATGCAAAAAAAAAAAATGAAAAAATGATTTCTTCGAAAAAA 874

Qy     841 attgtgactaataacatgactgactaaaaaaaaatctaaataa 879
           |||||||||||||||||||||||||||||||||||||||||||
Db     875 ATTGTGACTAATAACATGACTGACTAAAAAAAATCTAAATAA 915


RESULT   2
AX052773
LOCUS       AX052773                 879 bp    DNA     linear   PAT 13-JAN-2
DEFINITION  Sequence 5 from Patent WO0070045.
ACCESSION   AX052773
VERSION     AX052773.1  GI:12226945
KEYWORDS    .
SOURCE      Mus sp.
  ORGANISM  Mus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 879)
  AUTHORS   Jacobsson,A.T., Asadi,A.I., Westerberg,R.I., Nedergaard,J. and
            Tvrdik,P.M.
  TITLE     Fatty acid elongation genes and uses thereof
  JOURNAL   Patent: WO 0070045-A 23-NOV-2000;
            Karolinska Innovations AB (SE)
FEATURES             Location/Qualifiers
     source          1. .879
                     /organism="Mus sp."
                     /db_xref="taxon:10090"
BASE COUNT      194 a    267 c    274 g    144 t
ORIGIN


  Query Match             100.0%;  Score 879;  DB 6;  Length 879;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-254;
  Matches  879;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 atggaggcagctgaaggcctttgatagtgaagtgaatgctttcttggacaacatagttgtgca 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      37 ATGGAGGCAGCTGAAGGCCTTTGATAATGAAGTCAATGCTTTCTTGGACAACATAGTTGTGA 96

Qy      61 ccaagagattctcgagttcgagaggtagttcctactgaactcttacctcccacctttcatc 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      97 CCACGAGATTCTCGAGTTCGCGGGTGGTTCCTGCTGGACTCTTACCTTCCCACCTTCATC 156

Qy     121 ctcaccatcacgtacctgttctcgatatgactgctgtaacaagtacatgcaagaacaggcct 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     157 CTCACCATCACGTACCTGTTCTCGATATGACTGCTGTAACAAGTACATGCAAGAACAGGCCT 216

Qy     181 gctctgtctctcagagggcatcctcaccttgtataacctcgcaatcacactttttgtgcag 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     217 GCTCTGTCTCTCAGAGGGCATCCTCACCTTGTATAACCTCGCAATCACACTTTTTGTGCAG 276

Qy     241 tatatgctggtggaggctgtacctctccagctggagaaagaggttacaactttgcagtatcag 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     277 TATATGCTGGTGGAGCTCATCCTCTCCAGCTGGGAAAGAGGTTACAACTTTGCAGTGTCAG 336

Qy     301 aatctcgacaagtgcaagaagaaaatcaatgtcgagtacgcaagctcttgtgattgatgact 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     337 AATCTCGACAGTGCAGGAGAAGCTCAATGTCCGCGTACGCAAGCTCTTGTGCTGCTGACTAC 396

Qy     361 ttttccaaactagtggaattcttgaacacgatttttctttgtttctacgaaaaaaaaccaat 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     397 TTCTCCAAACTAGTGGAGTTCCTGAACACGATTTTTCTTTGTTTCTACGAAAAAAGACCAAT 456

Qy     421 cgaatcgacttcttcatgatgtatgtcgcgctcatgttcaacatctggtggtgtgtttt 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     457 CGAATCACCTTCCTTCATGTCTATCACCACGCGTCCATGTTCAACATCTGGTGGTGTGTTT 516

Qy     481 ttgaactggatacctttgtggtcaaagcttctttggacccaccctgaacagctttatccac 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     517 TTGAACTGGATACCTTTGTGGTCAAAGCTTCTTTGGACCCACCCTGAACAGCTTTATCCAC 576

Qy     541 attctcatgtactctactagcgctgtctgtgttcagtccatgcaaaagtacctttgg 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     577 ATTCTCATGTACTCTACTAGCGCTGTCTGTGTTCAGTCCATGCAAAAGTACCTTTGG 634

Qy     601 tggaaagaagtaactcctcaagcagctgaactgatgcagttcctactcaccatcatcgtcgct 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     635 TGGAAGAAGTACCTCACACAGCCTCAGTTGGTGCTGTTCGTACTCACCATCATCGTCGCT 694

Qy     661 ctgaatgccctgatgtgaggcctatgcttcctttgctgtctctcatttcgatgtttcctgt 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     695 CTGAGTGCCGTGGTGAAGGCCTCTGGCTTCCCCTTGCTGTCTATTTTCAGTTTCCTGTT 754

Qy     721 tatatgatgcagcgtgatcatcttttgaagcttctatattcgaagtatgcttcaaaaaaa 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     755 TATATGATGAGCGTGATCATCTTCTTGAAGCTTCTATATTCGAAGTATGCTTCAAAAAAA 814

Qy     781 cgaatgaaaaaaaaaatgaatgcaaaaaaaaaaaaatgaaaaaatgatttcttcgaaaaaa 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     815 CGAATGAAAAAAAAATGAATGCAAAAAAAAAAAATGAAAAAATGATTTCTTCGAAAAAA 874

Qy     841 attgtgactaataacatgactgactaaaaaaaaatctaaataa 879
           |||||||||||||||||||||||||||||||||||||||||||
Db     875 ATTGTGACTAATAACATGACTGACTAAAAAAAATCTAAATAA 915
```